

Matches	125;	Conservative	63;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	GACUAVUGUACUCUUAUAAAUAUUGAGUUAUUUCGAGUGUUAUAAAUAUUGUUAUUAUUG	60						
Db	1	GACTATGTATCTTTATTTAAATTAAGGTTAAATTTGACGGTTAAAAATGTTTAAATATG	60						
Qy	61	CUAUVGUCUUAAGGUCUGUUAUUAUUAUACUACACACAAAGUAGGACCGGAGACGCC	120						
Db	61	CTATAGTCTTAGAGGCTTGTAATATTTATTTACCAACAAAGATGACCGGAGACGCC	120						
Qy	121	UCCAAUUAUCUAGUGUACCCUGUGUCUGCUCUCAAACUUAAGUGUGUGUGCGAAAAA	180						
Db	121	TCCAATATCTAGTAGTACCTCGTGCTCGCTCAAAATTAAGTGAGTGTTGTGCGAAAAA	180						
Qy	181	UCUCACU	188						
Db	181	TCACACT	188						
RESULT 2									
LOCUS	BD177015	188 bp	RNA	linear	part 16-Apr-2003				
DEFINITION	Novel translational activity-promoting higher-order structure.								
ACCESSION	BD177015								
VERSION	BD177015.1	GI:30014275							
KEYWORDS	JP 2002306168-A/1.								
SOURCE	Plautia stali intestine virus								
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Dicistroviridae; Crispavirinae.								
REFERENCE	1 (bases 1 to 188)								
AUTHORS	Nakashima,N. and Kanamori,Y.								
TITLE	Novel translational activity-promoting higher-order structure								
JOURNAL	Patent: JP 2002306168-A 1 22-Oct-2002; DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND ENTOMOLOGICAL, HIROSHIMA PREFECTURAL INDUSTRIAL TECHNOLOGY PROMOTION ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND FISHERIES								
COMMENT	OS Plautia stali intestine virus PN JP 2002306168-A/1 PD 22-OCT-2002 PF 25-JAN-2001 JP 2001016746 PI NOBUHIKO NAKASHIMA,YASUSHI KANAMORI PC C12N5/09,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P21/02// CC (C12N15/09,C12R1/92),C12N15/00,C12N5/00,(C12N15/00,C12R1/92) CC Novel translational activity-promoting higher-order structure FH Key Location/Qualifiers FT source 1..188 FT /organism='Plautia stali intestine virus'. FEATURES source Location/Qualifiers 1..188 /organism="Plautia stali intestine virus" /mol_type="genomic RNA" /db_xref="taxon:64698"								
ORIGIN									
Query Match	100.0%;	Score 188;	DB 6;	Length 188;					
Match Local Similarity	66.5%;	Pred. No. 2e-34;							
Matches	125;	Conservative	63;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	GACUAVUGUACUCUUAUAAAUAUUGAGUUAUUUCGAGUGUUAUAAAUAUUGUUAUUAUUG	60						
Db	1	GACTATGTATCTTTATTTAAATTAAGGTTAAATTTGACGGTTAAAAATGTTTAAATATG	60						
Qy	61	CUAUVGUCUUAAGGUCUGUUAUUAUUAUACUACACACAAAGUAGGACCGGAGACGCC	120						
Db	61	CTATAGTCTTAGAGGCTTGTAATATTTATTTACCAACAAAGATGACCGGAGACGCC	120						
Qy	121	UCCAAUUAUCUAGUGUACCCUGUGUCUGCUCUCAAACUUAAGUGUGUGUGCGAAAAA	180						
Db	121	TCCAATATCTAGTAGTACCTCGTGCTCGCTCAAAATTAAGTGAGTGTTGTGCGAAAAA	180						
Qy	181	UCUCACU	188						

[illegible]

CDS

LENSKEICLHPQSWDSHIESFKHCAKLGSFPLFLDREPARSGVLNENY"
 <6473..9097
 /note="translation is supposed to start at several
 upstream codon, directed by internal ribosome entry site
 (IRIS)"
 /codon_start=1
 /product="capsid protein precursor"
 /protein_id="BA032554.1"
 /db_xref="GI:3493359"

/translation="ANNNNNNNNNSQKNDTTSDRENDPSVSGRIDSEVEFNOET
 HFADNAPVDSISLAGEITNLPKSLVTDHNRQHSVISFLQRPOLIKTVENAPGTAQGS
 LTTTIDIPDMLTSMVYDKLDGFATFPAADTIFRYQVNAQPOQCRILWAIYIPMDISLS
 TRTAEIRADIRIALPHVOLDISEQSEFTLRVPIYSYSAYNLEGRYGRVAVV
 YSPNOVSQPNLKVINIGYVNTNLTGTLTALSPAVARBOVINSEMDMRILAE
 SRNPPTKIAASINVIQKSDIIGVLPQTSPTNPVAKISDAFIDLSMTPEKRD
 KTNHGETVLPFRPTFGVNDGVESHKLGTHAMKRIDOPDFASKDEMSFDVVKAI
 LMYIDSFSYSNSNYGDTLWSTVSPCRSADYTTNGANRSPPTSTLYAIGPSS
 ISFTVPAAGTTPFKRIRPEINPLASSGISVDNFVATGVGVALPPLVIGSTVPS
 TIOIIVEMKGGPDEVEGPNSTGMPHISTIPATNGDTVSELVSTAQEGANFASG
 OHDRSDYLEDKIEIKDITGSSNISLATEKSLSCVGSRENPFDILKRCWPKQSV
 AFNTKILUSGIPVINYTSINGTGLTADGSTPLTNVSMWAFRGGFPKAYIHD
 LPABEMVQALINDSONTNPQPLALOSLOYLESDKLYEFSPWYCPYTLTTPSGS
 LNYISDLVNPPTTARITITISEYATAYAMAADDFCGFYIGAPLSMWEIERLAGRLD
 SSYGFVTSPLRVDPFKV"

ORIGIN

Query Match 34.3%; Score 64.4; DB 13; Length 9275;
 Best Local Similarity 39.8%; Pred. No. 2.1e-05;
 Matches 74; Conservative 43; Mismatches 66; Indels 3; Gaps 1;

OY 2 ACUAGUGAUCUUUUAUAUUGAGUUAAUUCGAGUUAAAAUAGUUUAUAVUGC 61
 DB 6289 AATGCTGATCTGATTAGAAAGTAG--AAAATCTTCAATTATATTTTAATACTGC 6345
 OY 62 UAUUGUCUUAAGGUCUUUUUAUUAUCUACACACAAGAGACCGAGACCCU 121
 DB 6346 TACATTTTAAAGACCTTAAAGTATTAGCTTACCCGACGAGATGGGGTTCCT 6405
 OY 122 CCAUUAUCUAGUUAUCCUUGUCUGUCUCAAACAUUAAGUGUGUUGUGGAAU 181
 DB 6406 GCAATATCCAGGCGACTAGGTGACGCTTGAATTAGTGACCTTAAGCTAAAGAT 6465
 OY 182 CUCACU 187
 DB 6466 TTCACT 6471

RESULT 8
 AB183472 9275 bp RNA linear VRL 13-JUL-2004
 LOCUS H1metcb1 P virus genomic RNA, complete sequence, isolate: izumo.
 DEFINITION AB183472
 VERSION AB183472.1 GI:50251148
 KEYWORDS
 SOURCE H1metcb1 P virus
 ORGANISM H1metcb1 P virus
 Virus; ssRNA positive-strand viruses, no DNA stage;
 Dicistroviridae; Crispavirus.

REFERENCE
 AUTHORS Nakashima, N., Kawahara, N., Omura, T. and Noda, H.
 TITLE Spherical viruses isolated from the brown planthopper, Nilaparvata
 lutea

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 9275)
 AUTHORS Nakashima, N. and Noda, H.
 TITLE Direct Submision
 JOURNAL Submitted (08-JUL-2004) Nobuhiko Nakashima, National Institute of
 Agrobiological Sciences, Otsu-shi 1-2, Tsukuba, Ibaraki 305-8634,
 Japan (E-mail:nakaj1@affrc.go.jp, Tel:81-29-838-6109,
 Fax:81-29-838-6028)

FEATURES
 source Location/Qualifiers
 1..9275

/organism="H1metcb1 P virus"
 /vifion
 /mol_type="genomic RNA"
 /isolate="izumo"
 /db_xref="taxon:81583"
 /lab_host="Nilaparvata lutea"
 /country="Japan:Shimane, Izumo"
 964..6297
 /codon_start=1
 /product="nonstructural protein precursor"
 /protein_id="BAD27584.1"
 /db_xref="GI:50251149"

CDS

/translation="MSOLIQTKISSYKNVYRKIKERSNINHPILRATPFVEDVEDSD
 LKINLEAIRTNVKNVVELLIGHQOLDGFCYLYNRDENLELVSNTSMNETHLPD
 LYNMLAEVYNNPNKQDILLGENSEMDQVYLSROLKEMRAIIVGQAGNVDILTPD
 DCQALAYTRVLYDAYVCRRARTYALCRILNHPHEIMPIYNDLIENFELCP
 FRRLKYLRCCLFSDMPLEIKWQCDERTEKIEIEVQDSVNVASQSDIE
 TLVSNMVEKSSVERMVAVCEBARQGFENVLAGINDTKGIDVLSIGVCS
 LYTDMLGLEVAASDILVNLKYLIVGLVAVALGVATGCVGMKSKLFNPLSVFSS
 DESMNNVAOQSSSDITWNGPLMLMSYFVSAGALIKSARCMNAIRCSMLPRAESGD
 TIFWIKSTYTMCRYIRSKYLVGDVPGVOVSADHPVASMLEEGEYKSPNGTSY
 DSATFSIHSLFIRGLALQGSBSPSQIARIGMDCLNKLITFERSNIEGVRNP
 PVVLYLHSGSGVKSILTNNVLASILSKIOPDMKLKQMKLITRASBDGFMDYTG
 QLVTFDDFQORDASGNPVVELDIVRANVYEPPLKMANLSOKASTNFTSKILICS
 SNLKQPKTESLNFENALYRRFDVCSYKNEKNDVPTFVEDFYQOEYDMLKED
 LGSTDMEGIVSKVELYKRSDFVSLDEKIOELTQSTOPEDEVPLNAQEQVNDVLG
 PCNDDCGEMTCWNTNHPKQWMLKMHVYTGIRKGSYVEAFKMRPLSOEYLA
 SKSRFGWLSIKERFPVIKDLRIHLVAVATVWNGPVFGVKKLPARKNECVLHVS
 ESIDVGNIKPRTESYTPVVKPTKITSYSPNVKSKTSSYESBNKPVREMSFPS
 GRVSGEMNIALSLQLDNVIKEVROQVSDNAAEICSLVTKMFKLYENDHVSILP
 GHVLFIKGRILAIIMPHEFLAALKFKPEKEGEGVVYFNNLFLRAPELIEDMIRVKKP
 ESPEPTESLAEBSRLGSCFLDNTINFDSGLFVSKDSLTKSSDILPPLSPNG
 QAPAKIRIGRAASGIDORDCRRVGSDDTDLRLVRYCYMRYLTERVEDGCAPIIARVY
 ALAKRIMGHIAENTAGSTPIYKEDDITILSMYLEQVANEQOQPIHMPGCHL
 PESTSFVVLKIEKPLVASSKSVISBPLHGLITPTKPEQLDPTDETSFQWRLK
 FASPCCVVDARMLENVSANVSNLCKSILENKDITLTSDSRSYFEAVSGIDEEFI
 NSVRSSSPGFVFDKEMNSKEKIFLEKEDVNTTKSLILRQOVEIISOAKLGV
 QOHVFIOTLDEREKIRAKTRMFASCPIDLYLACGMVFCGVVSLLOKSNIGISV
 GTNYSYSDWMTLIANTLISKSPCMTAGFEGEDSGOLDIILRAAGOVILNYSRDMGSG
 BEDLINVQVLESLSVSHLANNVYVMKGLSPGHELTALINISIPILSPSSWQIA
 FGVNKAFAFEFEVCGIYAGDDHIVPEPAINVFOYELASIFPKQIGLSTYLEDKO
 ATVNAAPRSINSEVYLKRFLEMDKQOYLAPLSEVLVLETPMNVKCCVDNLOTYTE
 LENSKEICLHPQSWDSHIESFKHCAKLGSFPLFLDREPARSGVLNENY"
 <6473..9097
 /note="translation starts at GCU (alanine) triplet,
 directed by internal ribosome entry site"
 /codon_start=1
 /product="capsid protein precursor"
 /protein_id="BAD27585.1"
 /db_xref="GI:50251150"

ORIGIN

Query Match 34.3%; Score 64.4; DB 13; Length 9275;
 Best Local Similarity 39.8%; Pred. No. 2.1e-05;
 Matches 74; Conservative 43; Mismatches 66; Indels 3; Gaps 1;

OY 2 ACUAGUGAUCUUUUAUAUUGAGUUAAUUCGAGUUAAAAUAGUUUAUAVUGC 61

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregiorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Huliyk,S., Hume,J., Idlebird,J., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpatly,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowalski,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshewa,L., Louiseged,H., Lorado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Mlilosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Muidasa,M., Murphy,M., Nair,L., Markovis,C., Neal,D., Morris,N., Nguyen,N., Norris,S., Nwokelamoh,O., Okwunu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Paeternak,S., Paul,H., Perez,A., Perez,L., Pflankoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,B., Pu,L.-L., Puzo,M., Quiroz,J., Rachlin,B., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Ritters,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shateman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Snaie,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Uemami,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 155949)

Worley,K.C.

Direct Submission

Submitted (24-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

REFERENCE

AUTHORS

TITLE

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

Center project name: KAYS
Center clone name: CH230-448E18
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 153428 bases at least Q40
Consensus quality: 154322 bases at least Q30
Consensus quality: 154784 bases at least Q20
Estimated insert size: 15661; sum-of-coverage estimation
Quality coverage: 8x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length
* (see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraftdata.html>).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 152687: contig of 152687 bp in length
* 152688 152787: gap of unknown length
* 152788 155949: contig of 3162 bp in length.
location/Qualifiers
1. 155949
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-448E18"
complement(1736..2514)
/note="clone boundary
clone_end:Sp6
site:
end_sequence:18251604"
complement(148932..149844)
/note="clone boundary
clone_end:T7
site:
end_sequence:BZ251602"
152688..152787
gap
/estimated_length=unknown
ORIGIN
Query Match 21.2%; Score 39.8; DB 14; Length 155949;
Best Local Similarity 49.4%; Pred. No. 8.6;
Matches 41; Conservative 15; Mismatches 27; Indels 0; Gaps 0;
QY 40 TTTAAATGAGTCTTAAATGCTTAAAGACAGAGGCTGCGATTTTCAACGACATG 109152
DB 109093 TTTAAATGAGTCTTAAATGCTTAAAGACAGAGGCTGCGATTTTCAACGACATG 109152
QY 100 CAAAGGACCGGAGCGCCCTC 122
DB 109153 CAGCATGCAAGGAGCGGAGCTC 109175
RESULT 12
AC098490/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-15D5, WORKING DRAFT SEQUENCE, 3
unnumbered pieces.
AC098490
AC098490.5 GI:30521857
VERSION
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
REFERENCE
Muzny,D.Marie, Metzger,M.Lee, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amth,A., Anguiano,D.,
Anyalebechi,V., Ayagi,A., Ayodeji,M., Baca,B., Baden,H.,

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

REFERENCE	1	Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Blatwal, C., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N.N., Bunney, C., Burch, P., Burrill, K., Calderon, E., Cardenas, V., Carter, C., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cize, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Decker, D., Delgado, O., Denison, S., Deremo, C., Ding, Y., Din, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Garza, M., Georegois, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hune, U., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Kapachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kwois, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, T., Liu, J., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, J., Lotenshuwar, L., Louisaed, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinec, E., Mathew, S., McLeod, M.P., McNally, T.Z., Meenen, B., Milosavljevic, A., Miner, G., Mnja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokoleh, G., Okunolu, G., Olarunmugbon, A., Pal, S., Parks, K., Paeternack, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polndexter, A., Popovic, D., Primus, E., Fu, L.-L., Plazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Rellly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.U., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, A., Steidle, M., Strong, R., Sutton, A., Svatek, A., Taboi, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.
REFERENCE	2	Unpublished
REFERENCE	3	Unpublished
REFERENCE	4	Unpublished
REFERENCE	5	Unpublished
REFERENCE	6	Unpublished
REFERENCE	7	Unpublished
REFERENCE	8	Unpublished
REFERENCE	9	Unpublished
REFERENCE	10	Unpublished
REFERENCE	11	Unpublished
REFERENCE	12	Unpublished
REFERENCE	13	Unpublished
REFERENCE	14	Unpublished
REFERENCE	15	Unpublished
REFERENCE	16	Unpublished
REFERENCE	17	Unpublished
REFERENCE	18	Unpublished
REFERENCE	19	Unpublished
REFERENCE	20	Unpublished
REFERENCE	21	Unpublished
REFERENCE	22	Unpublished
REFERENCE	23	Unpublished
REFERENCE	24	Unpublished
REFERENCE	25	Unpublished
REFERENCE	26	Unpublished
REFERENCE	27	Unpublished
REFERENCE	28	Unpublished
REFERENCE	29	Unpublished
REFERENCE	30	Unpublished
REFERENCE	31	Unpublished
REFERENCE	32	Unpublished
REFERENCE	33	Unpublished
REFERENCE	34	Unpublished
REFERENCE	35	Unpublished
REFERENCE	36	Unpublished
REFERENCE	37	Unpublished
REFERENCE	38	Unpublished
REFERENCE	39	Unpublished
REFERENCE	40	Unpublished
REFERENCE	41	Unpublished
REFERENCE	42	Unpublished
REFERENCE	43	Unpublished
REFERENCE	44	Unpublished
REFERENCE	45	Unpublished
REFERENCE	46	Unpublished
REFERENCE	47	Unpublished
REFERENCE	48	Unpublished
REFERENCE	49	Unpublished
REFERENCE	50	Unpublished
REFERENCE	51	Unpublished
REFERENCE	52	Unpublished
REFERENCE	53	Unpublished
REFERENCE	54	Unpublished
REFERENCE	55	Unpublished
REFERENCE	56	Unpublished
REFERENCE	57	Unpublished
REFERENCE	58	Unpublished
REFERENCE	59	Unpublished
REFERENCE	60	Unpublished
REFERENCE	61	Unpublished
REFERENCE	62	Unpublished
REFERENCE	63	Unpublished
REFERENCE	64	Unpublished
REFERENCE	65	Unpublished
REFERENCE	66	Unpublished
REFERENCE	67	Unpublished
REFERENCE	68	Unpublished
REFERENCE	69	Unpublished
REFERENCE	70	Unpublished
REFERENCE	71	Unpublished
REFERENCE	72	Unpublished
REFERENCE	73	Unpublished
REFERENCE	74	Unpublished
REFERENCE	75	Unpublished
REFERENCE	76	Unpublished
REFERENCE	77	Unpublished
REFERENCE	78	Unpublished
REFERENCE	79	Unpublished
REFERENCE	80	Unpublished
REFERENCE	81	Unpublished
REFERENCE	82	Unpublished
REFERENCE	83	Unpublished
REFERENCE	84	Unpublished
REFERENCE	85	Unpublished
REFERENCE	86	Unpublished
REFERENCE	87	Unpublished
REFERENCE	88	Unpublished
REFERENCE	89	Unpublished
REFERENCE	90	Unpublished
REFERENCE	91	Unpublished
REFERENCE	92	Unpublished
REFERENCE	93	Unpublished
REFERENCE	94	Unpublished
REFERENCE	95	Unpublished
REFERENCE	96	Unpublished
REFERENCE	97	Unpublished
REFERENCE	98	Unpublished
REFERENCE	99	Unpublished
REFERENCE	100	Unpublished
REFERENCE	101	Unpublished
REFERENCE	102	Unpublished
REFERENCE	103	Unpublished
REFERENCE	104	Unpublished
REFERENCE	105	Unpublished
REFERENCE	106	Unpublished
REFERENCE	107	Unpublished
REFERENCE	108	Unpublished
REFERENCE	109	Unpublished
REFERENCE	110	Unpublished
REFERENCE	111	Unpublished
REFERENCE	112	Unpublished
REFERENCE	113	Unpublished
REFERENCE	114	Unpublished
REFERENCE	115	Unpublished
REFERENCE	116	Unpublished
REFERENCE	117	Unpublished
REFERENCE	118	Unpublished
REFERENCE	119	Unpublished

AP008214_055 5500001 5610000
AP008214_056 5600001 5710000
AP008214_057 5700001 5810000
AP008214_058 5800001 5910000
AP008214_059 5900001 6010000
AP008214_060 6000001 6110000
AP008214_061 6100001 6210000
AP008214_062 6200001 6310000
AP008214_063 6300001 6410000
AP008214_064 6400001 6510000
AP008214_065 6500001 6610000
AP008214_066 6600001 6710000
AP008214_067 6700001 6810000
AP008214_068 6800001 6910000
AP008214_069 6900001 7010000
AP008214_070 7000001 7110000
AP008214_071 7100001 7210000
AP008214_072 7200001 7310000
AP008214_073 7300001 7410000
AP008214_074 7400001 7510000
AP008214_075 7500001 7610000
AP008214_076 7600001 7710000
AP008214_077 7700001 7810000
AP008214_078 7800001 7910000
AP008214_079 7900001 8010000
AP008214_080 8000001 8110000
AP008214_081 8100001 8210000
AP008214_082 8200001 8310000
AP008214_083 8300001 8410000
AP008214_084 8400001 8510000
AP008214_085 8500001 8610000
AP008214_086 8600001 8710000
AP008214_087 8700001 8810000
AP008214_088 8800001 8910000
AP008214_089 8900001 9010000
AP008214_090 9000001 9110000
AP008214_091 9100001 9210000
AP008214_092 9200001 9310000
AP008214_093 9300001 9410000
AP008214_094 9400001 9510000
AP008214_095 9500001 9610000
AP008214_096 9600001 9710000
AP008214_097 9700001 9810000
AP008214_098 9800001 9910000
AP008214_099 9900001 10010000
AP008214_100 1000001 10110000
AP008214_101 10100001 10210000
AP008214_102 10200001 10310000
AP008214_103 10300001 10410000
AP008214_104 10400001 10510000
AP008214_105 10500001 10610000
AP008214_106 10600001 10710000
AP008214_107 10700001 10810000
AP008214_108 10800001 10910000
AP008214_109 10900001 11010000
AP008214_110 11000001 11110000
AP008214_111 11100001 11210000
AP008214_112 11200001 11310000
AP008214_113 11300001 11410000
AP008214_114 11400001 11510000
AP008214_115 11500001 11610000
AP008214_116 11600001 11710000
AP008214_117 11700001 11810000
AP008214_118 11800001 11910000
AP008214_119 11900001 12010000
AP008214_120 12000001 12110000
AP008214_121 12100001 12210000
AP008214_122 12200001 12310000
AP008214_123 12300001 12410000
AP008214_124 12400001 12510000
AP008214_125 12500001 12610000
AP008214_126 12600001 12710000
AP008214_127 12700001 12810000

AP008214_128 12800001 12910000
AP008214_129 12900001 13010000
AP008214_130 13000001 13110000
AP008214_131 13100001 13210000
AP008214_132 13200001 13310000
AP008214_133 13300001 13410000
AP008214_134 13400001 13510000
AP008214_135 13500001 13610000
AP008214_136 13600001 13710000
AP008214_137 13700001 13810000
AP008214_138 13800001 13910000
AP008214_139 13900001 14010000
AP008214_140 14000001 14110000
AP008214_141 14100001 14210000
AP008214_142 14200001 14310000
AP008214_143 14300001 14410000
AP008214_144 14400001 14510000
AP008214_145 14500001 14610000
AP008214_146 14600001 14710000
AP008214_147 14700001 14810000
AP008214_148 14800001 14910000
AP008214_149 14900001 15010000
AP008214_150 15000001 15110000
AP008214_151 15100001 15210000
AP008214_152 15200001 15310000
AP008214_153 15300001 15410000
AP008214_154 15400001 15510000
AP008214_155 15500001 15610000
AP008214_156 15600001 15710000
AP008214_157 15700001 15810000
AP008214_158 15800001 15910000
AP008214_159 15900001 16010000
AP008214_160 16000001 16110000
AP008214_161 16100001 16210000
AP008214_162 16200001 16310000
AP008214_163 16300001 16410000
AP008214_164 16400001 16510000
AP008214_165 16500001 16610000
AP008214_166 16600001 16710000
AP008214_167 16700001 16810000
AP008214_168 16800001 16910000
AP008214_169 16900001 17010000
AP008214_170 17000001 17110000
AP008214_171 17100001 17210000
AP008214_172 17200001 17310000
AP008214_173 17300001 17410000
AP008214_174 17400001 17510000
AP008214_175 17500001 17610000
AP008214_176 17600001 17710000
AP008214_177 17700001 17810000
AP008214_178 17800001 17910000
AP008214_179 17900001 18010000
AP008214_180 18000001 18110000
AP008214_181 18100001 18210000
AP008214_182 18200001 18310000
AP008214_183 18300001 18410000
AP008214_184 18400001 18510000
AP008214_185 18500001 18610000
AP008214_186 18600001 18710000
AP008214_187 18700001 18810000
AP008214_188 18800001 18910000
AP008214_189 18900001 19010000
AP008214_190 19000001 19110000
AP008214_191 19100001 19210000
AP008214_192 19200001 19310000
AP008214_193 19300001 19410000
AP008214_194 19400001 19510000
AP008214_195 19500001 19610000
AP008214_196 19600001 19710000

Query Match 20.4%; Score 38.4; DB 15; Length 110000;
Best Local Similarity 32.1%; Pred. No. 20;
Matches 54; Conservative 33; Mismatches 81; Indels 0; Gaps 0;

QY 21 AUAAGGUAUAUUUCGAGUUAUUUUUAUUGUUAUGUUAUGAGGUCUG 80
Db 66297 ATTTCCTTAATTTTAAAGTTAAATTTTAACTATTATTATTAAGGATACATCAATA 66356
QY 81 UAAUUAUUAUACCAACAAGUGGACGGAGGCGCCUCCAUAUUAUGUUAUCCU 140
Db 66357 TTATATTATCTTCCCGATCAAAAAGCGCAAGCAATGAATGATCAAAAACTTGTTC 66416
QY 141 CGGCGUGGUAACAUAUAGUGUGUUGUGCGAAAAAGAUUCUACU 188
Db 66417 TTTTGTGTCCTATTATTAAGGAAATTTTAAAAAGAAATTCATT 66464

RESULT	15
AL772250/c	
LOCUS	AL772250
DEFINITION	Zebrafish DNA sequence from clone CH21-160M22, complete sequence.
ACCESSION	AL772250
VERSION	ALT772250.12
KEYWORDS	HTG.
SOURCE	Danio rerio (zebrafish)
ORGANISM	Danio rerio

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em1, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep

beginning 'dr' were identified by the Recon repeat discovery system (Zhurong Bao and Sean Eddy, submitted), and those beginning 'atr' further information see http://www/Projects/D_verito/fishmask.shtml

CH201-160M22 is from a CHOR1-211 BAC library

RECTOR: PTABAC2.1.

FEATURES
SOURCE

ORIGIN

Query Match	20.4%	Score 38.4	DB 5	length 159758
Best Local Similarity	35.2%	Pred. No. 18		
Matches 45	Conservative 27	Mismatches 56	Indels 0	Gaps 0

QY	40	UUAAAUUUUUUUUUUUUGCUAAGUCUUAAGAAGUCUUUGUAUUUUUAUCUUAACCA	99
Db	120586	TTAAATATATTTGTAAAAATCTTACTGTCAATCAATTTGATGATGTATGAACCTCGA	120527
QY	100	CAGAUGGACCGGAGACGCCUCCAAUUUUCUAGUGUACCCUGUGUGCCUCAAACAUUA	159
Db	120526	CTRGATTAAGCTCAATGATATCTGTCAATTATTACTCACTCTCTGTGCTATTTCAAACTTA	120467
QY	160	AGUGGUGU	167
Db	120466	ATTAATGT	120459

Search completed: December 22, 2005, 11:33:36
Job time : 1508.51 secs

Db 61 UCCAAGUCUUAAGUCUUGUUAUUUUUAUACUUAACCAACAAGAGGACCGGAGAGAGCC 120
OY 121 UCCAAGUCUUAAGUCUUGUUAUUUUUAUACUUAACCAACAAGAGGACCGGAGAGAGAA 180
Db 121 UCCAAGUCUUAAGUCUUGUUAUUUUUAUACUUAACCAACAAGAGGACCGGAGAGAGAA 180
OY 181 UCUCAACUU 188
Db 181 UCUCAACUU 188

RESULT 3
AA150544
ID AA150544 standard; RNA; 200 BP.
AC AA150544,
XX
XX 19-DEC-2002 (first entry)
XX
DE Plautia stali intestine virus RNA sequence.
XX
KW CrPV-like virus; 89; higher-order structure; drug development;
KM drug production; translational activity-promoting function;
KM protein synthesis; structural analysis.
XX
OS Plautia stali intestine virus.
XX
PN WO200261080-A1.
XX
PD 08-AUG-2002.
XX
PF 31-JAN-2001; 2001WO-JP000641.
XX
PR 25-JAN-2001; 2001JP-00016746.
XX
PA (NAG-) NAT INST AGROBIOLOGICAL SCI.
XX
PI Nakashima N, Kanamori Y;
XX
XX WPI; 2002-627482/67.
XX
DR Translational activity-promoting higher-order structure of CrPV-like
PT viruses for protein translation when suitably initiated, useful in
PT synthesis of proteins and polypeptides of foreign species for application
PT in drugs.
XX
XX
XX Claim 1; Fig 1-2; 38pp; Japanese.
XX
PS The invention comprises seven RNA sequences (CrPV-like viruses) which
CC have a higher-order structure that sustains translational activity-
CC promoting function. The RNA sequences of the invention are useful in the
CC synthesis of proteins and polypeptides for application in developing and
CC producing drugs. The RNA sequences of the invention are also useful in
CC basic research of protein synthesis and structural analysis by the gene
CC recombinant technique. The present nucleotide sequence represents a Plautia stali
CC intestine virus RNA sequence of the invention
XX
SQ Sequence 200 BP; 66 A; 34 C; 37 G; 0 T; 63 U; 0 Other;
Query Match 100.0%; Score 188; DB 6; Length 200;
Best Local Similarity 100.0%; Pred. No. 4.2e-43;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACUAGUCUUAUUUUUAUUUUUAUUUUUAUUUUUAUUUUUAUUUUUAUUUUUAUUUUUAUU 60
Db 1 GACUAGUCUUAUUUUUAUUUUUAUUUUUAUUUUUAUUUUUAUUUUUAUUUUUAUUUUUAUU 60
OY 61 CUUAAGUCUUAAGUCUUGUUAUUUUUAUACUUAACCAACAAGAGGACCGGAGAGAGCC 120
Db 61 CUUAAGUCUUAAGUCUUGUUAUUUUUAUACUUAACCAACAAGAGGACCGGAGAGAGCC 120
OY 121 UCCAAGUCUUAAGUCUUGUUAUUUUUAUACUUAACCAACAAGAGGAGUUGUGCGAAAAAGAA 180

Db 121 UCCAAGUCUUAAGUCUUGUUAUUUUUAUACUUAACCAACAAGAGGAGUUGUGCGAAAAAGAA 180
OY 181 UCUCAACUU 188
Db 181 UCUCAACUU 188

RESULT 4
AA235832
ID AA235832 standard; DNA; 430 BP.
AC AA235832;
XX
XX 03-FEB-2000 (first entry)
XX
XX Plautia stali intestine virus translation promoting sequence.
XX
XX Plautia stali intestine virus translation; promoter; synthesis;
KW ss.
XX
OS Plautia stali intestine virus.
XX
PN JP11290084-A.
XX
PD 26-OCT-1999.
XX
PF 10-APR-1998; 98JP-00114428.
XX
PR 10-APR-1998; 98JP-00114428.
XX
PA (NORQ) NORINSUISANSO SANSI KONCHU.
XX
XX WPI; 2000-016983/02.
XX
XX A DNA promoting translation activity to a protein - and synthesis of a
PT protein efficiently from a protein gene by using the DNA.
PT
XX
XX Claim 4; Page 2; 5pp; Japanese.
XX
PS The present sequence represents a Plautia stali intestine virus (PSIV)
CC translation promoting sequence. Also described in the present invention
CC are: (a) a plasmid having a gene for synthesizing the objective protein
CC downstream of the above DNA base sequence; (b) a plasmid in which the
CC above DNA base sequence is inserted between the promoter sequence and the
CC translation starting point of the objective gene; (c) a transformant in
CC which the above plasmid is introduced to a host cell; a recombinant
CC baculovirus having a gene for synthesizing the objective protein
CC downstream of the above DNA base sequence; (d) a recombinant baculovirus
CC in which the above DNA base sequence is inserted between the promoter
CC sequence and the translation starting point of the objective gene; (e) a
CC cultured cell infected by the above recombinant baculovirus; and (f) a
CC method for synthesizing efficiently a protein from a protein gene
CC position downstream by taking a DNA corresponding to a base sequence
CC promoting translation activity in the gene base sequence of an insect RNA
CC virus into a plasmid. The translation activating DNA is used for
CC synthesizing the encoded protein
XX
SQ Sequence 430 BP; 131 A; 80 C; 93 G; 126 T; 0 U; 0 Other;
Query Match 100.0%; Score 188; DB 3; Length 430;
Best Local Similarity 66.5%; Pred. No. 4.9e-43;
Matches 125; Conservative 63; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACUAGUCUUAUUUUUAUUUUUAUUUUUAUUUUUAUUUUUAUUUUUAUUUUUAUUUUUAUU 60
Db 235 GACTATGTGATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 294
OY 61 CUUAAGUCUUAAGUCUUGUUAUUUUUAUACUUAACCAACAAGAGGACCGGAGAGAGCC 120
Db 295 CTATATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 354
OY 121 UCCAAGUCUUAAGUCUUGUUAUUUUUAUACUUAACCAACAAGAGGAGUUGUGCGAAAAAGAA 180

[illegible]

	Best Local Similarity	97.9%	Pred. No. 2.8e-41	Matches 184	Conservative 0	Mismatches 4	Indels 0	Gaps 0
QY	1	GACUAGUAGUCCUUAUAAAUVAGAGUAAAUVUUCGAGGUTUAAAAUAGUTUUAUAVUG	60					
Db	81	GACUAGUAGUCCUUAUAAAUVAGUUAUAAUUVUCGAGGUTUAAAAUAGUTUUAUAVUG	140					
QY	61	CUAUVGUCUUAAGAGGUCUUGUUAUUUUUAUCUACACAAAGUAGACCGGAGAGCCC	120					
Db	141	CUAUVGUCUUAAGAGGUCUUGUUAUUUUUAUCUACACAAAGUAGACCGGAGAGCCC	200					
QY	121	UCCAAUUAUUAUGUUAACCCUGGCGGCGGCUCAAAAUUAAGUGGUGTUGGCGAAAAAGA	180					
Db	201	UCCAAUUAUUAUGUUAACCCUGGCGGCGGCUCAAAAUUAAGUGGUGTUGGCGAAAAAGA	260					
QY	181	UCUCACUU	188					
Db	261	UCUCACUU	268					

[illegible]

AA60324
ID AA60324 standard; RNA; 9829 BP.
XX
AC AA60324;
XX
DT 27-AUG-2003 (first entry)
XX
DE Aphid lethal paralysis virus (ALPV) RNA.
XX
KM ALPV; Aphid lethal paralysis virus; homopterous insect; floriculture;
XX aphid; white fly; agriculture; gene therapy; ss.
XX
OS Aphid lethal paralysis virus.
XX
PN WO2003040177-A2.
XX
PD 15-MAY-2003.
XX
PF 25-OCT-2002; 2002WO-NL000682.
XX
PR 25-OCT-2001; 2001NL-01019225.
XX
PA (PLAN-) PLANT RES INT BV.
XX
PI Van Der Wijk F, Van Munster M;
XX
XX WPI; 2003-441523/41.
XX
PT Novel nucleic acid derived from aphid lethal paralysis virus, useful for
XX detecting ALPV-viruses, for preparing a monoclonal antibody against
XX antigen, in research concerning aphids and for controlling aphids.
PS Claim 1; Page 10-13; 17pp; English.
XX
CC The invention relates to a novel RNA derived from an ALPV (aphid lethal
XX paralysis virus). ALPV RNA is useful for controlling homopterous insects
XX such as aphids or white flies. It is useful in the research concerning
XX aphids and for controlling aphids and in the field of agriculture or
XX floriculture. It is also useful for detecting ALPV-viruses using PCR and
XX for preparing a monoclonal antibody against antigen which is synthesised
XX in vitro. The invention is also useful in gene therapy. The present
XX sequence is ALPV RNA
SQ Sequence 9829 BP; 3083 A; 1906 C; 1879 G; 0 T; 2961 U; 0 Other;
Query Match 19.8%; Score 37.2; DB 8; Length 9829;
Best Local Similarity 55.4%; Pred. No. 2.2;
Matches 92; Conservative 0; Mismatches 73; Indels 1; Gaps 1;
QY 1 GACUAGUGAGUCUUAUUAAUUGGUUAUUUCGAGGUAAAUAUUGUUUAUUUUG 60
DB 6635 GAUUAUUAUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUU 6694
QY 61 C-UUAGUCUUAAGAGUCUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUU 119
DB 6695 CAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUU 6754
QY 120 CUCCAUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUU 165
DB 6755 CCAGCAAAUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUU 6800
RESULT 12
ABK39937
ID ABK39937 standard; DNA; 11422 BP.
XX
AC ABK39937;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human chemically pretreated gene sequence #9 strand 2.
XX
KM Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;

KW cytosolic; ALDH6; CYP11A; CYP11B1; DPVD; EPHX2; OCIN; TXNRD1;
XX UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO200202806-A2.
XX
PD 10-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP007470.
XX
PR 30-JUN-2000; 2000DE-01032529.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-154757/20.
XX
PT New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
XX useful for detecting cytosine methylation state of genes associated with
XX pharmacogenomics and for therapy of diseases e.g. cancer.
PS Claim 1; SEQ ID NO 18; 24pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence at least 18
XX bases in length of a segment of the chemically pretreated DNA of gene
XX CC associated with pharmacogenomics according to one of the sequences of the
XX genes ALDH6 (NM 000693), CYP11A (NM 000781), CYP11B1 (NM 000497), CYP3A3
XX (NM 000776 and NM 017460), DPVD (NM 000110), EPHX2 (NM 001979), OCIN
XX (NM 002538), TXNRD1 (NM 003330), UGT8 (NM 003360), MRP (NM 004996,
XX NM 019900, NM 019901, NM 019902, NM 019862, NM 019898, NM 019899) and
XX their complementary sequences, or a sequence (S1) chosen from 87
XX sequences and their complements. The chemical pretreatment is bisulphite
XX treatment to convert cytosines (but not methyl-cytosines) into uracils.
XX Also included are an oligomer (II) in particular an oligonucleotide or a
XX peptide nucleic acid (PNA)-oligomer, comprising in each case at least one
XX base sequence having a length of 9 nucleotides which hybridises to or is
XX identical to a chemically pretreated DNA of genes associated with
XX pharmacogenomics and their complements, arranged in an array for
XX analysing diseases associated with the methylation state (CpG) and/or
XX detecting SNPs (single nucleotide polymorphisms) of the 87 sequences. The
XX oligomers may also be used as PCR primers. The set of 87 nucleic acids
XX and their complements is useful for diagnosis and therapy of solid
XX tumours and cancer. The present sequence represents one the 87 DNA
XX sequences or its complement. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 11422 BP; 3913 A; 59 C; 1856 G; 5594 T; 0 U; 0 Other;
Query Match 19.3%; Score 36.2; DB 6; Length 11422;
Best Local Similarity 28.4%; Pred. No. 4.4;
Matches 23; Conservative 30; Mismatches 28; Indels 0; Gaps 0;
QY 13 UUAUUAAAUUAGGUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUU 72
DB 2581 TGATTAAAGTTTGTGATTATTAATTAATTAATTAATTAATTAATTAAGATTAAAGTTTAA 2640
QY 73 AGGUCUGUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUU 93
DB 2641 AGGTATTTTATTTGTTTAACTT 2661
RESULT 13
AB132219
ID AB132219 standard; DNA; 11422 BP.
XX
AC AB132219;
XX
DT 26-MAR-2002 (first entry)

This Page Blank (uspia)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 07:31:33 ; Search time 1967.29 Seconds

(without alignments)
4471.109 Million cell updates/sec

Title: US-10-088-750C-1

Perfect score: 188
1 gacuaugugucuaaauaa.....gugcgaaagaauucacuu 188

Sequence: 1 gacuaugugucuaaauaa.....gugcgaaagaauucacuu 188

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41076325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_sac1.*
2: gb_sac2.*
3: gb_sac3.*
4: gb_hnc.*
5: gb_sac4.*
6: gb_sac5.*
7: gb_sac6.*
8: gb_sac7.*
9: gb_gsa1.*
10: gb_gsa2.*
11: gb_gsa3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38.4	20.4	381	10	CL403205
2	38.4	20.4	715	10	CG155911
3	38.4	20.4	930	10	CG050490
4	38.4	20.4	958	9	CC438918
5	37.8	20.1	964	10	C2951756
6	37.6	20.0	604	10	BX134773
7	37.6	20.0	1101	10	CNS00EVL
8	36.8	19.6	373	10	CL448388
9	36.8	19.6	448	7	CO224988
10	36.4	19.4	666	3	B2386666
11	36.4	19.4	1133	10	AG3191971
12	36.2	19.3	666	3	B2386666
13	36	19.1	533	1	AM784857
14	36	19.1	571	9	CE130055
15	36	19.1	673	9	CC896465
16	36	19.1	709	10	CG304003
17	36	19.1	764	9	CC409240
18	36	19.1	934	10	CG103022
19	36	19.1	972	9	CC409237
20	35.8	19.0	1242	8	DN702302
21	35.8	19.0	835	2	BE877807
22	35.8	19.0	840	7	CO126087

23	35.6	18.9	454	11	DR1589T	AL748738	Danio rer
24	35.6	18.9	571	9	AZ238106	AZ238106	RPCR-23-7
25	35.6	18.9	596	11	DR21N16T	AL735446	Danio rer
26	35.6	18.9	715	10	CG151659	CG151659	PUDFX71TB
27	35.6	18.9	720	10	CM181274	CM181274	104_596_1
28	35.6	18.9	747	10	CM003045	CM003045	ZMMLA000
29	35.6	18.9	808	10	CM003264	CM003264	ZMMLA000
30	35.6	18.9	1101	10	CNS008X5	AL0552546	Drosophila
31	35.4	18.8	414	3	B1347107	B1347107	376521 MA
32	35.4	18.8	824	10	DU058171	DU058171	96654 Tom
33	35.2	18.7	1582	10	AG321318	AG321318	Mus muscu
34	35	18.6	664	10	CL175763	CL175763	104_381_1
35	35	18.6	682	9	CE219548	CE219548	tigr-gss-
36	34.8	18.5	420	7	CK546326	CK546326	rsmb0.01
37	34.8	18.5	432	10	CG850022	CG850022	ZMMLB003
38	34.8	18.5	437	3	B1936278	B1936278	PLEST0A2
39	34.8	18.5	454	7	CO274615	CO274615	EX14637.1
40	34.8	18.5	455	11	DE047749	DE047749	Oryzias 1
41	34.8	18.5	466	5	BQ451601	BQ451601	PFEST0A0
42	34.8	18.5	500	11	CR490394	CR490394	mtb2-161I
43	34.8	18.5	617	7	CO333954	CO333954	EK308442
44	34.8	18.5	679	10	C2678828	C2678828	OM_Be023
45	34.8	18.5	679	10	AG172534	AG172534	Pan trogl

ALIGNMENTS

RESULT 1
CL403205 381 bp DNA linear GSS 05-MAR-2004
LOCUS ZMMLB00412P06f ZMMLB (HindIII) Zee may genome clone
DEFINITION ZMMLB00412P06 5', genomic survey sequence.

ACCESSION CL403205
VERSION CL403205.1 GI:45213390
KEYWORDS GSS.
SOURCE Zee may
ORGANISM Zee may

REFERENCE
AUTHORS Bharti,A.K., Young,S., Kaychok,S., Keizer,G., Bronzino,A.C.,
Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
1 (bases 1 to 381)
clade; Panicoideae; Andropogoneae; Zee.

JOURNAL
TITLE Sequencing of the maize genome at Poir (2003c)
COMMENT Unpublished (2003)
CONTACT: Bharti,A.K.
Dr.Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: T7
Class: BAC ends
High quality sequence start: 93.
Location/Qualifiers
1..381
/organism="Zee may"
/mol_type="genomic DNA"
/culivar="B73"
/db_xref="taxon:4577"
/clone="ZMMLB00412P06"
/lab_host="E. coli DH10B"
/clone_id="ZMMLB (HindIII)"
/note="Vector: pCUGI, Site_1: HindIII, Site_2: HindIII"

ORIGIN

Query Match 20.4%; Score 38.4; DB 10; Length 381;
Best Local Similarity 39.2%; Pred. No. 10;
Matches 47; Conservative 22; Mismatches 51; Indels 0; Gaps 0;

OY		19	AAATUAGGUAUAUUUCGAGGUAAAAAAGUTUUUAUUUGCUAUAGCUTCUAGAGGUCU	78
Db		153	AAATTAGGGGTATAACGAGGCGAAAATTATATTTTACATTTTTTATCATTTTAATATGA	212
OY		79	UGUAUUAUUUAUVCUUAACCAACAAGUGAGCCGAGACGCCCUCCAUAUUAUGUGUACC	138
Db		213	GTTTTTTTTTAAAAACGACTTAAGGGTATTAGGTAGACACACCACTAAACTTAAGGCACC	272
RESULT 2				
LOCUS	CG155911/c		715 bp	DNA linear GSS 21-AUG-2003
DEFINITION	PUMJ87TD ZM 0.6_1.0 KB zea mays genomic clone ZMMB7A0616006,			
ACCESSION	CG155911			
VERSION	CG155911.1			
KEYWORDS	GSS.			
SOURCE	Zea mays			
ORGANISM	Zea mays			
TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
JOURNAL	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD			
COMMENT	clade; Panicoidae; Andropogoneae; Zea. 1 (bases 1 to 715) WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Reenick,A., Frazer,C.M., Yuan,Y., San Miguel,P., Ma,U. and Bennerzen,J. Maize Genomics Consortium Unpublished (2003) Other GSSs: PUMJ87TB Contact: Cathy WhiteLaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whiteLaw@cigr.org Seq primer: TP Class: sheared ends. Location/Qualifiers 1..715 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMB7A0616006" /clone_1lb="ZM_0.6_1.0_KB" /note="Vector: PCR4-TOPo; site_1: EcoRI; 0.6-1.0 kb high Cor selected genomic DNA library"			
FEATURES				
source				
Query Match		20.4%	Score 38.4;	DB 10; Length 715;
Best Local Similarity		39.2%	Fred.No.10;	
Matches	47/	Conservative	22;	Mismatches 51; Indels 0; Gaps 0;
OY		19	AAAUUAGGUUAUUUCGAGGUAAAAAAGUTUUUAUUUGCUUAAGCUTCUAGAGGUCU	78
Db		588	AAATTAGGGGTATAACGAGGCGAAAATTATATTTTACATTTTATTCATTTTAATATGA	529
OY		79	UGUAUUAUUUAUVCUUAACCAACAAGAUGCAGCCGAGCACCCCTCCAUUAUUCUAGUACC	138
Db		528	GTTTTTTTTTAAAAACGACTTAAGGGTATTAGGTAGACACCACTAAACTTAAGGCACC	469
RESULT 3				
LOCUS	CG050490		930 bp	DNA linear GSS 19-AUG-2003
DEFINITION	PUDCX20TD ZM 0.6_1.0_KB zea mays genomic clone ZMMB7A0639D16,			
ACCESSION	CG050490			
VERSION	CG050490.1			
KEYWORDS	GSS.			
SOURCE	Zea mays			
ORGANISM	Zea mays			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			

[illegible]

Qy	Db	Query Match	Best Local Similarity	Matches	Score	Pred. No.	Mismatches	Indels	Gaps
Qy	Db	19	AAATUAGGUNAATUUCGAGGUTUAAAUAUAGUUAUUGUUAUAGUCUAGAGUCU	78	20.4%	39.2%	22	51	0
Qy	Db	66	AAATAGGTGTATTAACGAGGGAATTAATTAATTTATTTATTTATTAATGA	125	32.1%	39.2%	22	51	0
Qy	Db	79	UGUAUUUAUUAUCUACACACAAAGUAGCAGCAGCCUCCAAUUAUAGUAGUACC	138	20.4%	39.2%	22	51	0
Qy	Db	126	GTITTTTTTAAACACGACTTAAGGCTTATGATGACACACCAAACTTAAGGACCC	185	20.4%	39.2%	22	51	0
RESULT 5	LOCUS	CZ951756	664 bp	DNA	linear	GSS	11-AUG-2005		
DEFINITION	LOCUS	263042 Tomato EcorI BAC library Lycopersicon esculentum genomic							
ACCESSION	LOCUS	clone_SL_EcorI0053E09 5, genomic survey sequence.							
VERSION	LOCUS	CZ951756							
KEYWORDS	LOCUS	CZ951756.1 GI:72295630							
SOURCE	LOCUS	GSS.							
ORGANISM	LOCUS	Lycopersicon esculentum (Solanum lycopersicum)							
REFERENCE	LOCUS	Lycopersicon esculentum							
AUTHORS	LOCUS	Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.							
TITLE	LOCUS	1 (bases 1 to 964)							
JOURNAL	LOCUS	Mueller, L.A., Bueler, R.M., Wang, Y., Tankale, S.D., Giovannoni, J.J., Van Eck, J., and Stack, S.							
COMMENT	LOCUS	BAC end sequencing from three Solanum lycopersicon libraries unpublished (2005)							
	LOCUS	Other_GSSs: 267643							
	LOCUS	Contact: Lukas Mueller							
	LOCUS	Tankale Lab, Dept. of Plant Breeding							
	LOCUS	Cornell University							
	LOCUS	251 Emerson Hall, Ithaca, NY 14853, USA							
	LOCUS	Tel: 607-255-6557							
	LOCUS	Fax: 607-255-6583							
	LOCUS	Email: agn-feedback@cornell.edu							
	LOCUS	Plate: 53 row: E column: 9							
	LOCUS	Seq primer: T7							
	LOCUS	Class: BAC ends							
	LOCUS	High quality sequence start: 4							
	LOCUS	High quality sequence stop: 544.							
	LOCUS	Location/Qualifiers							
	LOCUS	1. .964							
	LOCUS	/organism="Lycopersicon esculentum"							
	LOCUS	/mol_type="genomic DNA"							
	LOCUS	/culivar="Heinz 1706"							
	LOCUS	/db_xref="taxon:4081"							
	LOCUS	/clone="SL_EcorI0053E09"							
	LOCUS	/lab_host="E. coli"							
	LOCUS	/clone_11b="Tomato EcorI BAC library"							
	LOCUS	/note="Vector: unk; Site_1: EcorI"							
ORIGIN									
Query Match		20.1%	Score 37.8;	DB 10;	Length 964;				
Best Local Similarity		34.0%	Pred. No. 15;						
Matches		33;	Conservative 27;	Mismatches 37;	Indels 0;	Gaps 0;			
9	GAUCUUAUAUAUAGGUAUAUUCGAGGUAUAUAUAGUUAUAGUUAUAGUUAUAGUC	68							
756	GTITTTACAAAATTAATTTATTTCTTAATTAATTAATTTAACTTTAACTCTTAATTT	815							

RESULT 6	EX134773/c	604 bp	DNA	linear	GSS 28-JAN-2003
LOCUS	EX134773	604 bp	DNA	linear	GSS 28-JAN-2003
DEFINITION	Danio rerio genomic clone DKEX-84C14, genomic survey sequence.				
ACCESSION	EX134773				
VERSION	EX134773.1	GI:27966066			
KEYWORDS	GSS.				
SOURCE	Danio rerio (zebrafish)				
ORGANISM	Danio rerio				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.				
AUTHORS	1 (bases 1 to 604)				
JOURNAL	Humphray, S.J., Huckle, E. and Durham, J.L.				
FEATURES	Direct Submission				
SOURCE	Submitted (27-JUN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Unpublished				
	This sequence was generated from the T7 end of BAC 84C14. 84C14 is part of the Daniokey BAC library created by R. Plasterk and N.V. Keygene. Further details: http://www.sanger.ac.uk/Projects/D_rerio/				
	Location/Qualifiers				
	1..604				
	/organism="Danio rerio"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:7955"				
	/clone="DKEX-84C14"				
	/tissue_type="Testis"				
	/note="vector pindigobAC-536"				
ORIGIN					
Query Match	20.0%;	Score 37.6;	DB 10;	Length 604;	
Best Local Similarity	38.0%;	Pred. No. 17;			
Matches	35;	Conservative 23;	Mismatches 34;	Indels 0;	Gaps 0;
QY	26	GUUAAUUCGAGGUGUAAAUAAGUUUUUAUUGCAUUGGUCUUAAGUCUUCUUAAGGUCUGUUUAU	85		
DB	319	GTTTACTGCAATGATAGAAAATTGGAATATTTCACCAAACTTATTGTTATTGTTAT	260		
QY	86	UUAUACUUAACACGACAGAUAGGACCGAGGACG	117		
DB	259	ATGTATTTTTAACATTAGAACTAGCTAAACAG	228		
RESULT 7					
CNS00EVL					
LOCUS	CNS00EVL	1101 bp	DNA	linear	GSS 04-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC: BACR23B23 of RPCR-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL069706				
VERSION	AL069706.1	GI:4949849			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1101)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila				

	Query Match Similarity	19.4%	Score 36.4	DB 10,	Length 1133;	
	Best Local Similarity	31.3%	Pred. No. 36;			
	Matches	52;	Conservative	33;	Mismatches 81;	Indels 0; Gaps 0
Oy		9	GAUCUUAUUAAAUAGGUAUAUUUCGAGGUUUAAAAUAGUUUUUAUAUUGCUAUAUGC	68		
Db		938	GAAUUTATTAATTAGAAGGGGCGAGTTTGGCGTTAAAATTATATATAGTTGGCCGGGC	997		
Oy		69	UUAAGGUCUUGUAUUUAUUUAUACCAACAAGAUAGGCCGAGCCGCCCAUAUU	128		
Db		998	TGGTCATAATTATATATTTGTATTATGAAGACGAGTGATCATATPATAGACTAAGATPA	1057		
Oy		129	CUAGGUACCCUCGUCGUCGCUCAAACUUAAGUGGUGUGCCGA	174		
Db		1058	COAGAAATATGTGGGTGGGAGACTATATAGATGTGTTATAGCGTGA	1103		
RESULT 12						
LOCUS	BJ382574	666 bp	mRNA	linear	EST 08-MAR-2000	
DEFINITION	BJ382574 Dictyostelium discoideum cDNA library, CF Dictyostelium					
ACCSSION	BJ382574					
VERSION	BJ382574.1	GI:19291958				
KEYWORDS	EST.					
SOURCE	Dictyostelium discoideum					
ORGANISM	Dictyostelium discoideum					
REFERENCE	Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.					
AUTHORS	Unruhshara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.					
TITLE	Full length cDNA of Dictyostelium discoideum at the culmination stage					
JOURNAL	unpublished (2002)					
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 111 Yata, Mishima, Shizuoka 411-8540, Japan Tel.: 81-559-81-6856 Fax: 81-559-81-6855 Email: teshin@genes.nig.ac.jp. Location/Qualifiers					
FEATURES						
source	1..666					
	/organism="Dictyostelium discoideum"					
	/mol_type="mRNA"					
	/strain="AX4"					
	/db_xref="taxon:44689"					
	/clone="ddc45n05"					
	/sex="mat A"					
	/dev_stage="Culmination stage"					
	/clone_lib="Dictyostelium discoideum cDNA library, CF"					
ORIGIN						
Query Match	19.3%	Score 36.2;	DB 3;	Length 666;		
Best Local Similarity	28.6%	Pred. No. 40;				
Matches	26;	Conservative	30;	Mismatches 35;	Indels 0; Gaps 0;	
Oy		4	UAUUGUAGUUAUUAAUUAUAGUUUAUUUCGAGUUAAAAUAGUUUUUAUAUUGCA	63		
Db		545	TAGTTGACACAATTAAAAATTAATTAATTAATTAATTTGTTTATTATTGA	604		
Oy		64	UAGUCUUAAGAGUCUUGUAUUUAUUAUCUA	94		
Db		605	AAATATTAAAAATTAATTAATTAATTAATTA	635		
RESULT 13						
LOCUS	AM784857	533 bp	mRNA	linear	EST 09-JUL-2000	
DEFINITION	11155 MARC 1PIG Sue scrofa cDNA 5', mRNA sequence.					
ACCSSION	AM784857					
VERSION	AM784857.1	GI:7841633				
KEYWORDS	EST.					
SOURCE	Sue scrofa (pig)					

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Sus scrofa	1 (bases 1 to 533)	Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J., Valler,J., Wise,T., Rohrer,G.A., Perlee,G., Sultana,R., Quackenbush,J. and Keefe,J.W.	Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly	Mamm. Genome 13 (8), 475-478 (2002)	
Mamm. Genome 13 (8), 475-478 (2002)	12226715	Contact: Smith TPL	USDA, ARS, US Meat Animal Research Center	PO Box 166, Clay Center, NE 68933-0166, USA	
Tel: 402 762 4366	Fax: 402 762 4390	Email: smtlh@mail.marc.usda.gov	Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.		
PCR primers	FORWARD: AGGAACAGCTATGACCAT	BACKWARD: GTTTCCTCAGTCACGACG	Plate: 45	row: D	column: 6
Seq primer: ATTTAGGTGACACTATG.	Location/Qualifiers	1..533	/organism="Sus scrofa"	/mol_type="mRNA"	/db_xref="taxon:9823"
			/db_xref="taxon:9823"	/cbase_type="pooled"	/lab_host="DH108"
			/clone_id="MARC 1P1G"	/note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI; library made from pooled tissue from day 11, 15, 15, 20, and 30 embryos."	
ORIGIN	Query Match	19.1%; Score 36; DB 1; Length 533;	Best Local Similarity	38.1%; Pred. 45;	Matches
	Matches	32; Conservative	22; Mismatches	30; Indels	0; Gaps
QY	22	UUAGUUAAUUCGAGUUAAAUAUUGUUUAUUAUUGCUUAUUGUUAGUUUUU	81		
Db	262	TTAAATTTAAATTTATTTGAGCTTAAAGTTGATTATTAATCAATTAAGCTTTAAAGACCTGG	321		
QY	82	AUAUUUUAUACUUAACACACAAGAU	105		
Db	322	ACATCTAGACTTCAGCAAAACAT	345		
RESULT 14	CE130055	571 bp	DNA	linear	GSS 25-SEP-2003
LOCUS	CE130055/c	ttgtr-gss-dog-17000326068677	Dog Library	Canis familiaris genomic,	
DEFINITION	genomic survey sequence.				
VERSION	CE130055				
ACCESSION	CE130055.1	GI:35228961			
KEYWORDS	GSS.				
SOURCE	Canis familiaris (dog)				
ORGANISM	Canis familiaris				
	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;				
REFERENCE	1 (bases 1 to 571)				
AUTHORS	Kirkness,E.F., Batina,V., Halpern,A.L., Levy,S., Remington,K., Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and Venter,J.C.				
TITLE	The dog genome: survey sequencing and comparative analysis				
JOURNAL	Science 301 (5641), 1898-1903 (2003)				
PubMed	14512627				
COMMENT	Contact: Kirkness EF				

131 AGUGUACCCUCGUGCUCGCCAAACAUVAGUGUGUUUGCGAAAAGAUC 182
| : : : | : : : | : : :
239 ATGCATATTTCATGACTTATACATGAATACTGCTGGTTTGTAGATTTC 290

this Page Blank (uspic)

Result No.	Score	Query Match	Length	DB	ID	Description
1	34.8	18.5	759	3	US-09-270-767-5116	Sequence 5136, App
2	34.8	18.5	759	3	US-09-270-767-20418	Sequence 20418, A
3	33.6	17.9	54945	3	US-09-967-669-10	Sequence 10, Appl
4	33.6	17.9	65485	3	US-09-949-016-13757	Sequence 13757, A
5	33.6	17.9	65552	3	US-09-949-016-12303	Sequence 12303, A
6	33.2	17.7	601	3	US-09-949-016-29048	Sequence 29048, A
7	33.2	17.7	601	3	US-09-949-016-68590	Sequence 68590, A
8	33	17.6	360470	3	US-09-949-016-13173	Sequence 13173, A
9	32.4	17.2	1338	3	US-09-830-330A-96	Sequence 96, Appl
10	32.4	17.2	1410	3	US-09-830-330A-95	Sequence 95, Appl
11	32.2	17.1	1345	3	US-09-270-767-1072	Sequence 1072, Appl
12	32.2	17.1	1345	3	US-09-270-767-16354	Sequence 16354, A
13	32.2	17.1	96922	3	US-09-949-016-17061	Sequence 17061, A
14	31.8	16.9	11349	3	US-09-949-016-16932	Sequence 16932, A
15	31.8	16.9	80595	3	US-09-078-294-3	Sequence 3, Appl
16	31.8	16.9	524032	3	US-09-949-016-16928	Sequence 16928, A
17	31.8	16.9	524032	3	US-09-949-016-16929	Sequence 16929, A
18	31.8	16.9	524032	3	US-09-949-016-16930	Sequence 16930, A
19	31.8	16.9	524032	3	US-09-949-016-16931	Sequence 16931, A
20	31.8	16.9	529885	3	US-09-949-016-14340	Sequence 14340, A
21	31.8	16.9	529885	3	US-09-949-016-14341	Sequence 14341, A
22	31.8	16.9	529885	3	US-09-949-016-14342	Sequence 14342, A
23	31.8	16.9	529885	3	US-09-949-016-14343	Sequence 14343, A
24	31.8	16.9	529885	3	US-09-949-016-14344	Sequence 14344, A

C 25	31.8	16.9	529885	3	US-09-949-016-13345	Sequence 14345, A
C 26	31.8	16.9	529885	3	US-09-949-016-13345	Sequence 14346, A
C 27	31.8	16.9	529885	3	US-09-949-016-13347	Sequence 14347, A
C 28	31.6	16.8	529807	2	US-08-935-466-3	Sequence 3, Appl1
C 29	31.4	16.7	23167	3	US-09-949-016-16836	Sequence 16836, B
C 30	31.4	16.7	50383	3	US-09-949-016-11600	Sequence 17600, A
C 31	31.4	16.7	53558	3	US-09-949-016-16616	Sequence 16616, A
C 32	31.2	16.6	113366	3	US-09-949-016-12277	Sequence 12277, A
C 33	31.2	16.6	113367	3	US-09-949-016-17051	Sequence 17051, A
C 34	31.2	16.5	113367	3	US-09-949-016-14780	Sequence 14780, A
C 35	30.8	16.5	7261	3	US-09-9270-767-10838	Sequence 10838, A
C 36	30.8	16.4	22750	3	US-09-949-016-17175	Sequence 17175, A
C 37	30.6	16.3	7183	3	US-09-081-148-9	Sequence 9, Appl1
C 38	30.6	16.3	7183	3	US-09-081-149-10	Sequence 10, Appl1
C 39	30.6	16.3	8093	3	US-09-949-016-15241	Sequence 15241, A
C 40	30.6	16.3	9048	3	US-08-973-272-4	Sequence 4, Appl1
C 41	30.6	16.3	10320	3	US-09-949-016-11778	Sequence 11778, A
C 42	30.6	16.3	11485	3	US-09-410-464-9	Sequence 9, Appl1
C 43	30.6	16.3	46805	3	US-09-949-002-585	Sequence 585, App
C 44	30.6	16.3	46806	3	US-09-949-002-842	Sequence 842, App
C 45	30.6	16.3	251672	3	US-09-949-016-17296	Sequence 17296, A

ALIGNMENTS

```

RESULT 1
US-09-270-767-5136
; Sequence 5136, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5136
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-5136

Query Match      18.5%; Score 34.8; DB 3; Length 759;
Best Local Similarity 37.2%; Pred. No. 0.29;
Matches 29; Conservative 22; Mismatches 27; Indels 0; Gaps 0;

QY
5 AUGUGAUCUUAUUAAUUGUUAAUUUUCGAGUUAAAAUUGUUUAUUGCUAU 64
| | | | | : : : : : | : : : : : | : : : : : | : : : : : |
Db 615 AAGCCATTGATTTAAATTATGCTAAATTTAAATCTAAAAATAATCTTAATTGCATT 674

QY
65 AGUCUUGAGGUCUUGUA 82
| | | | | : : : : :
Db 675 TTGCCCATAGTATTTTAA 692

RESULT 2
US-09-270-767-20418
; Sequence 20418, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20418
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Drosophila melanogaster

```



```

; ORGANISM: Human
US-09-949-016-17061
Query Match      17.1%; Score 32.2; DB 3; Length 96922;
Best Local Similarity 23.7%; Pred. No. 9;
Matches 22; Conservative 33; Mismatches 38; Indels 0; Gaps 0;
QY      2 ACUAGUCGUAUUUAAUUGUUAUUUUCGAGUAAAUAUUGUUUAUUUUGC 61
DB      28021 ATTATCAAAATGCTCTAAAGCTTTAAAGTCTTTTAAATAATGTTTATATCTT 28080
QY      62 UAUAGUCUUGAGGUCUCUUAUUUAUUAUUAUUA 94
DB      28081 TATAATTTGAATAATTTTATGTTTCTCTGTTA 28113

RESULT 14
US-09-949-016-16932
; Sequence 16932, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16932
; LENGTH: 11349
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16932
Query Match      16.9%; Score 31.8; DB 3; Length 11349;
Best Local Similarity 32.5%; Pred. No. 6.1;
Matches 27; Conservative 24; Mismatches 32; Indels 0; Gaps 0;
QY      1 GACUAGUGAUCUUAUUUAAUUGUUAUUUUCGAGUUAUUUUAUUUAUUUAUUG 60
DB      8810 GATAGTTTGAACCTTAAACACGATTTCTTTTGTGTAATAATTGAGTATAAACATTTG 8869
QY      61 CUUAGUCUUAUGAGUCUCUUAU 83
DB      8870 ATAGTTTCTTAGAGTTGTTTAT 8892

RESULT 15
US-09-078-294-3
; Sequence 3, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 80595
; TYPE: DNA
; ORGANISM: Nucleotide sequence of HC-conf1g
US-09-078-294-3
```

```

Query Match      16.9%; Score 31.8; DB 3; Length 80595;
Best Local Similarity 24.1%; Pred. No. 11;
Matches 20; Conservative 31; Mismatches 32; Indels 0; Gaps 0;
QY      12 CUUUAUUAAAUAUAGGUUAAUUVCGAGUUAAAUAUUGUUUAUUUUGCUAUAUCUUA 71
DB      392 CTTTTTAAATTAATGATTTTATGATATAATTTTCCTTTTATTTTAAAGTTTAA 451
QY      72 GAGUCUCUUAUUUAUUAUUA 94
DB      452 GATGTATGTTTCCTTAACTTA 474

Search completed: December 22, 2005, 15:30:10
Job time : 86.0668 secs
```


THIS PAGE BLANK (USPTO)


```

; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 192
; LENGTH: 11422
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-192

Query Match      19.3%; Score 36.2; DB 6; Length 11422;
Best Local Similarity 28.4%; Pred. No. 31;
Matches 23; Conservative 30; Mismatches 28; Indels 0; Gaps 0;

QY      13 UUAUUAUUAGUUAUUUCAGGUAUUUUUUAUUAGUUUAUAUUUGCUUAUAGUCUUAAG 72
DB      2581 TGATTAAAGTTTGTATTGATATTAATTAATTAATTAATTAATTAATTAAGTTTAA 2640
        |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
        |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY      73 AGGUCUUGUAUUUAUUAUCU 93
DB      2641 AGGTATTTATGTGTTTAGTT 2661
        |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 6
US-10-257-166-18
; Sequence 18, Application US/10257166
; Publication No. US20040023230A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPERBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of
; TITLE OF INVENTION: Genes Implicated in Pharmacogenomics
; FILE REFERENCE: 5013.1011
; CURRENT APPLICATION NUMBER: US/10/257,166
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: PCT/EP01/07470
; DE 10032329.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-06-29
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 18
; LENGTH: 11422
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-257-166-18

Query Match      19.3%; Score 36.2; DB 7; Length 11422;
Best Local Similarity 28.4%; Pred. No. 31;
Matches 23; Conservative 30; Mismatches 28; Indels 0; Gaps 0;

QY      13 UUAUUAUUAGUUAUUUCAGGUAUUUUUUAUUAGUUUAUAUUUGCUUAUAGUCUUAAG 72
DB      2581 TGATTAAAGTTTGTATTGATATTAATTAATTAATTAATTAATTAATTAAGTTTAA 2640
        |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
        |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY      73 AGGUCUUGUAUUUAUUAUCU 93
DB      2641 AGGTATTTATGTGTTTAGTT 2661
        |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 7
US-10-017-161-743/c
; Sequence 743, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
```

```

; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 743
; LENGTH: 1099
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1099)
; NAME/KEY: CDS
; LOCATION: (201)..(556)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (587)..(899)
US-10-017-161-743

Query Match      19.0%; Score 35.8; DB 6; Length 1099;
Best Local Similarity 32.3%; Pred. No. 17;
Matches 54; Conservative 31; Mismatches 82; Indels 0; Gaps 0;

QY      20 AAUAGGUUAUUUUCAGGUUAAAAUAGUUUUUAUUAGCUUAUAGUCUUAAGGUCU 79
DB      967 ACTAAAGTTAAATACAGTCTCTCAATAAATAGAAAGATTTTATTATTAATTAAGATA 908
        :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
        :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY      80 GUUAUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 139
DB      907 CACTGATTTTACTCAAAAGACGTTCTTCTGATCAGTTTCCAAAAGCTGTTTAC 848
        :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
        :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY      140 UCGUCUCGCUCAACAUAUUAUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 186
DB      847 TCCTGTCTCAAACTATGATGATGCGGTTCAACATAGGGCTTAC 801
        :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
        :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 8
US-10-292-798-649/c
; Sequence 649, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 649
; LENGTH: 1099
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1099)
; NAME/KEY: CDS
; LOCATION: (201)..(556)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (587)..(899)
```


US-10-292-798-649

```
Query Match      19.0%; Score 35.8; DB 6; Length 1099;
Best Local Similarity 32.3%; Pred. No. 17;
Matches 54; Conservative 31; Mismatches 82; Indels 0; Gaps 0;

QY 20 AAUAGUGUAAUUCGAGGUUAAAUAUAGUUUAUUGCUUAUAGUCUUAAGUCU 79
DB 967 ACRAAGTTAAATACGATCTCTCAATTAAGAAATTTTATTGACTAATTAAGATA 908
QY 80 GUAAUUAUUAUCUACACACAAAGAGGACCCGAGACCCUCCAUUAUCUAGUACCC 139
DB 907 CACTGATTACTTCAAAAAGACGTTCTTCTGATCAGTTCCAAAAGCTTGTTTACA 848
QY 140 UCGUGCTCGCACAACUUAUAGUGUCUUCGCAAAAGAUCCAC 186
DB 847 TCCTTGCTCTCAAACTATAGATGATGGGTTCAACTAGGCTTTAC 801
```

RESULT 9

```
US-09-925-065A-440970/c
; Sequence 440970, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252, 147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250, 092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261, 766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289, 846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 440970
; LENGTH: 577
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-440970
```

```
Query Match      18.9%; Score 35.6; DB 4; Length 577;
Best Local Similarity 32.2%; Pred. No. 15;
Matches 29; Conservative 27; Mismatches 34; Indels 0; Gaps 0;
```

```
QY 1 GACUAGUGUACUUAUUAUUAUAGGUAAUUCGAGGUUAAAUAUAGUUUAUUAUG 60
DB 449 GAATTAATTTTACTTTTAAATTAAGTAAATGATGATTTCTTAAGAGCTTTAAGCTTG 390
QY 61 CUAAUGUCUUAAGAGUCUCUUAUUAUUA 90
DB 389 CTAAATTCAGTAATAATTATTAATA 360
```

RESULT 10

```
US-10-027-632-160930/c
; Sequence 160930, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
```

```
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146, 002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160930
; LENGTH: 908
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-160930
```

```
Query Match      18.9%; Score 35.6; DB 5; Length 908;
Best Local Similarity 34.0%; Pred. No. 18;
Matches 36; Conservative 26; Mismatches 44; Indels 0; Gaps 0;
```

```
QY 1 GACUAGUGUACUUAUUAUUAUAGGUUAAAUAUAGGUUAAAUAUAGUUUAUUAUG 60
DB 840 GAAGATGTTTGGTATTAATTAATTAAGTCTACAGAGGCTTAAGCATGTTATTTT 781
QY 61 CUAAUGUCUUAAGAGUCUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 106
DB 780 AAAAGTATCTCTCTTGTAGATTATTAATTAATTAATTAATTAATTAATTAATTA 735
```

RESULT 11

```
US-10-027-632-160931/c
; Sequence 160931, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146, 002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160931
; LENGTH: 908
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-160931
```

```
Query Match      18.9%; Score 35.6; DB 5; Length 908;
Best Local Similarity 34.0%; Pred. No. 18;
Matches 36; Conservative 26; Mismatches 44; Indels 0; Gaps 0;
```

```
QY 1 GACUAGUGUACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 60
```


RESULT 15
US-09-938-842A-5166
; Sequence 5166, Application US/09938842A


```

; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58759
; LENGTH: 2828
; TYPE: DNA
; ORGANISM: Bovine 19866880519533
US-10-750-185-58759
```

```
Query Match      17.6%; Score 33; DB 6; Length 2828;
Best Local Similarity 30.7%; Pred. No. 1.9;
Matches 47; Conservative 31; Mismatches 75; Indels 0; Gaps 0;
```

```
QY 2 ACUATGUCUUAUUUUAAGGUAUUUUGAGGUAUUUUUAGUUUUUAUUUUAUUUUC 61
DB 2409 ACAATTCATCTGTAGATCACTTAATTTAGAGAGAGAAACCTTAATGATTTAA 2468
QY 62 UAUAGUCUUAAGGUCUGUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUA 121
DB 2469 TATATGAGGTGGGGTCTTAAGTATCTTAAATCCAGCTGTGATCAATTAACAGCT 2528
QY 122 CCAUUAUUUAAGUUAUCCUUGUCUGUCUCUCAA 154
DB 2529 GTGATTAAGAGTCACTACATAGCTGTGGCTCAGA 2561
```

```
RESULT 3
US-11-121-086-78
; Sequence 78, Application US/11121086
; Publication No. US20050265459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 78
; LENGTH: 189993
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-78
```

```
Query Match      17.1%; Score 32.2; DB 7; Length 189993;
Best Local Similarity 29.4%; Pred. No. 13;
Matches 25; Conservative 27; Mismatches 33; Indels 0; Gaps 0;
```

```
QY 10 AUGCUUAUUUAUUUUAAGGUAUUUUGAGGUAUUUUUAUUUUAUUUUAUUUUAUUUUA 69
DB 138808 ATATATATAAAATTAATTAAGTATATTATATATATATGAATTTATTTATTTT 138867
QY 70 UAGAAGUCUUGUAUUUUAUUUUAUUUUA 94
DB 138868 GAAATATATATATATATATATATATATA 138892
```

```
RESULT 4
US-10-750-185-20455/C
; Sequence 20455, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
```

```

; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20455
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MM8T02818
US-10-750-185-20455
```

```
Query Match      16.8%; Score 31.6; DB 6; Length 600;
Best Local Similarity 32.6%; Pred. No. 2.9;
Matches 28; Conservative 24; Mismatches 34; Indels 0; Gaps 0;
```

```
QY 19 AAUUAAGUUAUUUUGAGGUAUUUUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUA 78
DB 496 AAATTTATTTAACTCTAGAGTATCTATTTGTTTCAAATAGTTTATTTTGTGTG 437
QY 79 UGUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUA 104
DB 436 TTGTTTCTGAAAGTACCATATATAA 411
```

```
RESULT 5
US-10-750-185-44380/C
; Sequence 44380, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44380
; LENGTH: 1330
; TYPE: DNA
; ORGANISM: Bovine 19866880991643
US-10-750-185-44380
```

```
Query Match      16.8%; Score 31.6; DB 6; Length 1330;
Best Local Similarity 32.6%; Pred. No. 3.8;
Matches 28; Conservative 24; Mismatches 34; Indels 0; Gaps 0;
```

```
QY 19 AAUUAAGUUAUUUUGAGGUAUUUUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUA 78
DB 1113 AAATTTATTTATTTCTAGAGTATCTATTTGTTTCAATAGTTTATTTATTTTGTGT 1054
QY 79 UGUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUAUUUUA 104
DB 1053 TTGTTTCTGAAAGTACCATATATAA 1028
```


	Query Match	16.8%	Score 31.6;	DB 6;	Length 1917;
	Best Local Similarity	28.2%;	Pred. No. 4.3;		
	Matches	31;	Conservative	30;	Mismatches 49; Indels 0; Gaps 0;
Oy	12	CUUUAUUAAAUAGGCUAAAUUGCGAGGUAAAAUUAUUAUUAUUGCUAAGUCUUA	71		
		:::::	:::::		
Db	1629	CCTTATCATATTGGTTGAGGTTCATACCAACAATTTTATTBAACATTTCTTTTGTGTT	1688		
		:::::	:::::		
Oy	72	GAGGUCUGUAUUUUUAUACUUAACCAACAAGAUGACCAGCAGCCGU	121		
		:::::	:::::		
Db	1689	TAGTTTTTCATTTTTTAACATATTACAAAGAAATGCTCTTGGAACATCT	1738		
		:::::	:::::		

Query Match	16.7%	Score 31.4	DB 6	length 600
Best Local Similarity	35.2%	Pred. No. 3.4		
Matches 32; Conservative	22	Mismatches 37	Indels 0	Gaps 0

QY 14 TAATUAAAUAAGGUAUAUUUUCGAGGUAUAAUAUGUUUUAUUUUGCUAAUAGUCUUAAG 73
 DB 349 TAATTAATAATGCTTTAAAGTAAACATTATTAATAATGGCAGTGTATTACTCTWAGAGTTTAAAG 290

```

: RESULT 8
: US-10-857-780-6/c
: Sequence 6, Application US/10857780
: Publication No. US20050272043A1
: GENERAL INFORMATION:
: APPLICANT: ROTH, RICHARD B.
: APPLICANT: BRAUN, ANDREAS
: APPLICANT: KAMMERER, STEFAN M.
: APPLICANT: NELSON, MATTHEW ROBERTS
: APPLICANT: RENELAND, RIKARD HENRY
: APPLICANT: HOYAL-WRIGHTSON, CAROLYN R.
: TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
: TITLE OF INVENTION: THEREOF
: FILE REFERENCE: SEQ.4063-CP
: CURRENT APPLICATION NUMBER: US/10/857,780
: CURRENT FILING DATE: 2004-05-28
: PRIOR APPLICATION NUMBER: 10/723,681
: PRIOR FILING DATE: 2003-11-25
: PRIOR APPLICATION NUMBER: 60/490,234
: PRIOR FILING DATE: 2003-07-24
: PRIOR APPLICATION NUMBER: 60/525,239
: PRIOR FILING DATE: 2003-11-25
: NUMBER OF SEQ. ID NOS: 4962
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 6

```

```

1 TYPE: DNA
2 ORGANISM: Homo sapiens
3 FEATURE:
4 NAME/KEY: misc feature
5 LOCATION: (3185)..(3315)
6 OTHER INFORMATION: n is a, c, g, or t
7 FEATURE:
8 NAME/KEY: misc feature
9 LOCATION: (7923)..(7932)
10 OTHER INFORMATION: n is a, c, g, or t
11 FEATURE:
12 NAME/KEY: misc feature
13 LOCATION: (10272)..(10272)
14 OTHER INFORMATION: n is a, c, g, or t
15 FEATURE:
16 NAME/KEY: misc feature
17 LOCATION: (28873)..(28873)
18 OTHER INFORMATION: n is a, c, g, or t
19 FEATURE:
20 NAME/KEY: misc feature
21 LOCATION: (29105)..(29106)
22 OTHER INFORMATION: n is a, c, g, or t
23 FEATURE:
24 NAME/KEY: misc feature
25 LOCATION: (35901)..(35901)
26 OTHER INFORMATION: n is a, c, g, or t
27 FEATURE:
28 NAME/KEY: misc feature
29 LOCATION: (36488)..(36488)
30 OTHER INFORMATION: n is a, c, g, or t
31 FEATURE:
32 NAME/KEY: misc feature
33 LOCATION: (40260)..(40260)
34 OTHER INFORMATION: n is a, c, g, or t
35 FEATURE:
36 NAME/KEY: misc feature
37 LOCATION: (52594)..(52594)
38 OTHER INFORMATION: n is a, c, g, or t
39 FEATURE:
40 NAME/KEY: misc feature
41 LOCATION: (68230)..(68230)
42 OTHER INFORMATION: n is a, c, g, or t

```



```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (170027)..(70027)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (83514)..(83514)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (89477)..(89477)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (108762)..(108762)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (109378)..(109378)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (113639)..(113639)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (115312)..(115312)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (115315)..(115315)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (127817)..(127817)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (138224)..(138224)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140476)..(140476)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147488)..(147488)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147500)..(147500)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147527)..(147527)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147534)..(147534)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (171395)..(171395)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (195657)..(195657)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (196415)..(196415)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:

```

```

NAME/KEY: misc_feature
LOCATION: (196417)..(196417)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (21236)..(21236)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (212442)..(212442)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (212445)..(212445)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (212455)..(212455)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (215163)..(215163)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (217657)..(217657)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (217725)..(217725)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (231757)..(231757)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (231762)..(231762)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (235827)..(235827)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (241414)..(241414)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (248915)..(248915)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (250679)..(250679)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (250575)..(250575)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature

```

```

Query Match      16.7%; Score 31.4; DB 6; length 285300;
Best Local Similarity 31.9%; Pred. No. 25;
Matches 36; Conservative 26; Mismatches 51; Indels 0; Gaps 0;

QY      16 UUAUAAUAGGUUAUUUUGAGGUUAUUUUAUUUUUAUUUUGCUUUAUUGCUUUAAGG 75
Db      48938 TTGTAGATATTTTATATATAGACTTTTCAATTTTAAATTTTGTGATACATAGT 48879
QY      76 UCUUUAUUUUUAUUUACUACACACAGAGUGAGCCGAGCCCUCAUUAU 128
Db      48878 AGATGTAGATATTTTATGCGGTACATGATGATGTTTGTGATACAGCATGATGT 48826

```



```
RESULT 9
US-10-750-185-41557/c
; Sequence 41557, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Denise
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41557
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Bovine 19866881267522
US-10-750-185-41557

Query Match      16.5%; Score 31; DB 6; Length 909;
Best Local Similarity 34.2%; Pred. No. 5;
Matches 27; Conservative 22; Mismatches 30; Indels 0; Gaps 0;

Oy      10 AUCUUAUAAAUAUAGCUAAUUCGAGUAAAUAUUCUUAUUAUUCUUAUAGUCU 69
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      730 ATATTAATGAGAAATGTTTAATTGATGTGAAATCAAGTTGAAATGATGTGTT 671
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      70 UAGAGUCUUGUAUUAUUA 88
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      670 CATGGCTTGTGTATATGA 652
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-11-058-727-15
; Sequence 15, Application US/11058727
; Publication No. US20050261483A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Prensall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/287809
; CURRENT APPLICATION NUMBER: US/11/058,727
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (truncated)
; FEATURE:
; FEATURE:
```

```
NAME/KEY: CDS
LOCATION: (10)...(1860)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (0)...(0)
OTHER INFORMATION: 49PVD
US-11-058-727-15

Query Match      16.4%; Score 30.8; DB 7; Length 1860;
Best Local Similarity 33.1%; Pred. No. 7.3;
Matches 43; Conservative 25; Mismatches 62; Indels 0; Gaps 0;

Oy      14 UAUUAAAUAUAGUAAUUCGAGUAAAUAUAGUUUAUUAUUCUUAUAGUCUUA 73
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1008 TAGTAATCTTCACAAATGTATGAACTAATCAAAATCTACAGCACTAGACTTGA 1067
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      74 GGUUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 133
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1068 TTTTACGAATTATGATTTTATCAAGACTATCAAGAGATGCAAGTACTCTTGATATTGT 1127
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      134 GUACCCUUCU 143
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1128 TTACCTCTGT 1137
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-11-108-389-15
; Sequence 15, Application US/11108389
; Publication No. US20050261188A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Prensall
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins with Pesticidal
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 35718/291049
; CURRENT APPLICATION NUMBER: US/11/108,389
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 10/606,320
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis (truncated)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10)...(1860)
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: 49PVD
US-11-108-389-15

Query Match      16.4%; Score 30.8; DB 7; Length 1860;
Best Local Similarity 33.1%; Pred. No. 7.3;
Matches 43; Conservative 25; Mismatches 62; Indels 0; Gaps 0;

Oy      14 UAUUAAAUAUAGUAAUUCGAGUAAAUAUAGUUUAUUAUUCUUAUAGUCUUA 73
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1008 TAGTAATCTTCACAAATGTATGAACTAATCAAAATCTACAGCACTAGACTTGA 1067
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      134 GUACCCUUCU 143
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1128 TTACCTCTGT 1137
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```



```

? PRIOR APPLICATION NUMBER: 10/606,320
? PRIOR FILING DATE: 2003-06-25
? NUMBER OF SEQ ID NOS: 134
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 11
? LENGTH: 2010
? TYPE: DNA
? ORGANISM: Bacillus thuringiensis (truncated)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(2010)
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (0)..(0)
? OTHER INFORMATION: 1218-1A
US-11-058-727-11

```

Query Match	16.4%	Score	30.8	DB	7	Length	2010
Best Local Similarity	33.1%	Pred. No.	7.5				
Matches	43	Conservative	25	Mismatches	62	Indels	0
						Gaps	0

[illegible]

```

RESULT 15
US-11-108-389-11
Sequence 11, Application US/11108389
Publication No. US20050261189A1
GENERAL INFORMATION:
APPLICANT: Andre R. Abad
APPLICANT: Ronald D. Flammagan
APPLICANT: Rafael Herrmann
APPLICANT: Theodore W. Kahn
APPLICANT: Albert L. Lu
APPLICANT: Billy Fred McCutchen
APPLICANT: James K. Prensall
APPLICANT: James F.H. Wong
APPLICANT: Cao-Guo Yu
TITLE OF INVENTION: Genes Encoding Proteins With Peptidicidal
TITLE OF INVENTION: Activity
FILE REFERENCE: 35718/291049
CURRENT APPLICATION NUMBER: US/11/108,389
CURRENT FILING DATE: 2005-04-18
PRIOR APPLICATION NUMBER: 60/351,786
PRIOR FILING DATE: 2002-06-26
PRIOR APPLICATION NUMBER: 60/460,787
PRIOR FILING DATE: 2003-04-04
PRIOR APPLICATION NUMBER: 10/606,320
PRIOR FILING DATE: 2003-06-25
NUMBER OF SEQ ID NOS: 134
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 2010
TYPE: DNA
ORGANISM: Bacillus thuringiensis (truncated)
FEATURES:
NAME/KEY: CDS
LOCATION: (1) ... (2010)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (0)...(0)
OTHER INFORMATION: 1218-1A
US-11-108-389-11

```

Query Match	16.4%;	Score 30.8;	DB 7;	Length 2010;
Best Local Similarity	33.1%;	Pred. No. 7.5;		
Matches	43;	Conservative	25;	Mismatches 62;
				Indels 0;
				Gaps 0;

[illegible]

Search completed: December 22, 2005, 16:40:56
Job time : 178.55 secs

This Page Blank (usps.gov)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 04:46:32 ; Search time 1475.19 Seconds
(without alignments)
7205.664 Million cell updates/sec

Title: US-10-088-750C-2

Perfect score: 187
Sequence: 1 gaaagaugugauguacuaa.....uaggcuagaagaauucacua 187

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_in: *
3: gb_env: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pr: *
9: gb_ro: *
10: gb_atb: *
11: gb_gy: *
12: gb_un: *
13: gb_vl: *
14: gb_mtg: *
15: gb_pl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	187	100.0	187 6	BD173512 Novel ter
2	187	100.0	187 6	BD177016 Novel ter
3	187	100.0	9275 13	AB017037 Himetobi
4	187	100.0	9275 13	AB183472 Himetobi
5	64.4	34.4	188 6	BD173511 Novel ter
6	64.4	34.4	188 6	BD177015 Novel ter
7	64.4	34.4	430 6	E28980 DNA promoti
8	64.4	34.4	8797 13	AB006531 Plautia s
9	64.4	34.4	188 6	BD173514 Novel ter
10	46.4	24.8	188 6	BD177018 Novel ter
11	46.4	24.8	9185 13	AF218039 Cricket P
12	40.2	21.5	160056 14	AC161224 Mus muscu
13	40.2	21.5	212284 14	AC131732 Mus muscu
14	40.2	21.5	229783 9	AC157478 Mus muscu
15	40	21.4	164647 5	CR846749 Danio rer
16	38.4	20.5	159143 5	CR352297 Zebrafish
17	38.4	20.5	232064 14	AC137218 Rattus no
18	38.4	20.5	252150 14	AC098201 Rattus no

19	38	20.3	174607 9	AC137555 Mus muscu
20	38	20.3	218575 8	AC132603 Mus muscu
21	37.4	20.0	156572 8	AC104108 Homo sapi
22	37.2	19.9	73190 8	AC130416 Homo sapi
23	37.2	19.9	161757 14	AC044901 Homo sapi
24	37.2	19.9	180058 9	AC133175 Mus muscu
25	37.2	19.9	192541 9	AC153502 Mus muscu
26	37.2	19.9	253591 14	AC163476 Bos tauru
27	37	19.8	190 6	BD173516 Novel ter
28	37	19.8	190 6	BD177020 Novel ter
29	37	19.8	8550 13	AF183905 Black que
30	36.8	19.7	225119 14	AC097781 Rattus no
31	36.8	19.7	232555 14	AC128470 Rattus no
32	36.6	19.6	179580 14	AC160586 Atelesrix
33	36.6	19.6	195117 14	AC149872 Xenopus t
34	36.2	19.4	181109 8	AF002456 Homo sapi
35	36.2	19.4	150193 14	AC009641 Homo sapi
36	36.2	19.4	175518 5	BX537270 Zebrafish
37	36.2	19.4	182064 14	BX537163 Danio rer
38	36	19.3	189 6	BD173513 Novel ter
39	36	19.3	189 6	BD177017 Novel ter
40	36	19.3	1345 6	AR496112 Sequence
41	36	19.3	1345 6	AR511394 Sequence
42	36	19.3	9264 13	AF014388 Drosophila
43	36	19.3	80309 15	AC022288 Arabidops
44	36	19.3	125632 15	AC015446 Arabidops
45	36	19.3	154071 2	AC115598 Dictyoste

ALIGNMENTS

RESULT 1	BD173512	187 bp	RNA	linear	PAT 18-FEB-2003
LOCUS	BD173512				
DEFINITION	Novel tertiary structure having ability to accelerate translation activity.				
ACCESSION	BD173512.1	GI:28414843			
KEYWORDS	WO 02061080-A/2.				
SOURCE	Himetobi P virus				
ORGANISM	Himetobi P virus				
REFERENCE	1 (bases 1 to 187)				
AUTHORS	Nakashima,N. and Kanamori,Y.				
TITLE	Novel tertiary structure having ability to accelerate translation activity				
JOURNAL	Patent: WO 02061080-A 2 08-AUG-2002;				
	JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF				
	STRUCTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE				
	FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA,YASUSHI KANAMORI				
COMMENT	OS Himetobi P virus				
	PN WO 02061080-A/2				
	PD 08-AUG-2002				
	PF 31-JAN-2001 WO 2001JP000641				
	PR 25-JAN-2001 JP 01P 016746				
	PI NOBUHIKO NAKASHIMA,YASUSHI KANAMORI				
	PC C12N15/11,C12N15/86,C12P21/02				
	CC Novel tertiary structure having ability to accelerate translation activity				
	FH Key				
	FT source				
FEATURES	Location/Qualifiers				
source	1..187				
	/organism="Himetobi P virus"				
	/mol_type="genomic RNA"				
ORIGIN	/db_xref="taxon:81583"				
Query Match	100.0%; Score 187; DB 6; Length 187;				
Best Local Similarity	66.3%; Pred. No. 1e-36;				

Matches	124;	Conservative	63;	Mismatches	0;	Indels	0;	Gaps	0;
Oy	1	GAAAAGUGUGAUCGUAUUAAGAAAGAAAUAUCCUGAUAUAUAUUUUUAUAUCGC 60							
Db	1	GAAAATGCTGATCTGATTAGTAAGAAATAATCTCTAGTATATATTTTAAATATCTGC 60							
Oy	61	UACAUUUUAAGACCCUUAAGUUUAUUAAGCCGCCAGAGUGGUGAGCGUUCU 120							
Db	61	TACATTTTAAAGACCCCTTAGTATTAGCTTTACGCCAGAGATGGGTGACGCTTCT 120							
Oy	121	GCAAUUACGAGGCGACCUAGGUGGACCCUUGUUAUUAUGUGACUUUAGCCUAAAGAU 180							
Db	121	GCAATATCCAGGCGACCTTAGGTGACGCTTGATGTTTAGTGAGCTTAAGCTAAAGAT 180							
Oy	181	UUCACUA 187							
Db	181	TTCACTA 187							

RESULT 2
BD177016 187 bp RNA linear PAT 16-APR-2003
LOCUS BD177016 Novel translational actively-promoting higher-order structure.
DEFINITION BD177016.1 GI:30014276
ACCESSION JP 2002306168-A/2.
VERSION Himetobi P virus
KEYWORDS Himetobi P virus
SOURCE Himetobi P virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Dicitroviridae; Cripavirus.

REFERENCE
AUTHORS 1 (bases 1 to 187)
TITLE Nakashima,N. and Kanamori,Y.
JOURNAL Novel translational activity-promoting higher-order structure
Parent: JP 2002306168-A 2 22-OCT-2002;
DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND
ENTOMOLOGICAL HIROSHIMA PREFECTURAL INDUSTRIAL TECHNOLOGY PROMOTION
ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND
FISHERIES

COMMENT
OS Himetobi P virus
PN JP 2002306168-A/2
PD 22-OCT-2002
PF 25-JAN-2001 JP 2001016746

FEATURES
source 1.187
Location/Qualifiers
/organism="Himetobi P virus"
/mol_type="genomic RNA"
/db_xref="taxon:81583"

ORIGIN
Query Match 100.0%; Score 187; DB 6; Length 187;
Best Local Similarity 66.3%; Pred. No. 1e-36;
Matches 124; Conservative 63; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GAAAAGUGUGAUCGUAUUAAGAAAGAAAUAUCCUGAUAUAUAUUUUUAUAUCGC 60
Db 1 GAAAATGCTGATCTGATTAGTAAGAAATAATCTCTAGTATATATTTTAAATATCTGC 60
Oy 61 UACAUUUUAAGACCCUUAAGUUUAUUAAGCCGCCAGAGUGGUGAGCGUUCU 120
Db 61 TACATTTTAAAGACCCCTTAGTATTAGCTTTACGCCAGAGATGGGTGACGCTTCT 120
Oy 121 GCAAUUACGAGGCGACCUAGGUGGACCCUUGUUAUUAUGUGACUUUAGCCUAAAGAU 180
Db 121 GCAATATCCAGGCGACCTTAGGTGACGCTTGATGTTTAGTGAGCTTAAGCTAAAGAT 180
Oy 181 UUCACUA 187

CDS

Db 181 TTCACTA 187

RESULT 3
AB017037 9275 bp RNA linear VRL 22-JAN-2000
LOCUS AB017037 Himetobi P virus genomic RNA, complete sequence.
DEFINITION AB017037.1 GI:3493357
ACCESSION AB017037
VERSION nonstructural protein precursor; capsid protein precursor.
KEYWORDS Himetobi P virus
SOURCE Himetobi P virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Dicitroviridae; Cripavirus.

REFERENCE
AUTHORS 1 (sites)
TITLE Nakashima,N., Sasaki,J. and Toriyama,S.
Determining the nucleotide sequence and capsid-coding region of
himetobi P virus: a member of a novel group of RNA viruses that
infect insects

JOURNAL Arch. Virol. 144 (10), 2051-2058 (1999)
PUBMED 10550677
REFERENCE 2 (bases 1 to 9275)
AUTHORS Nakashima,N. and Sasaki,J.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-1998) Nobuhiko Nakashima, National Institute of
Sericultural and Entomological Science, Department of Insect
Physiology and Behavior, 1-2 Owashi, Tsukuba, Ibaraki 305-8634,
Japan (E-mail:naka@nises.affrc.go.jp, Tel:81-298-38-6109,
Fax:81-298-38-6028)

FEATURES
source 1.9275
Location/Qualifiers
/organism="Himetobi P virus"
/mol_type="genomic RNA"
/specific_host="Lacodelphax striatellus"
/db_xref="taxon:81583"
964 . 6297
/codon_start=1
/product="nonstructural protein precursor"
/protein_id="BA332553.1"
/db_xref="GI:3493358"

/translation="MSQLIQTKISYKNYVNOIKKRGSRNHPILRATPNFVEDVESD
LKNLEAIRITTFVFNKVVLELTGHODLFQFYKYNRLLENLVYMTSMNETHLHPD
LVNLEIVANNPNKDLDLCEGNSMPOQYLSQLEKMRALVWGLQAGVDFLIDTVNP
DCOIALYMYTRVLVYAKVCSRRVRARVALGLINNVPHLIPYINKDLINPELCPV
PRKXVWVYLTKRCILAESDPLEIRWQDVCDERMKEKIEBQEQVYVNSRSRSPTE
TVLSNREYCSGNSVERKVAVCCEARQGLPVRVSLGINDITGYGIDSVSRIGNGVCS
LITDTMLGLEVAASDLVNLKTYIGFLVAVLVGATYGVGMISKLFNPLSLVKS
DESMNVAOESGSDTWNQPLMLMSVFGVSAGALKSAKSNMIRICSMPLPAESGID
TIFTWIKSTYTCYRIFSKYVLGVDPGVVSADSHFVASMLELGEFPGYSFNGTFSY
DSATFSLHSFLTRGLNQRSFSPSDQAIRGMDCKNKLITFSPSRNIEGVSANP
PVVIVHGGSGVSKTLTNVLAASLTKLOPMNKKOONKLYSASQDEFMDGVTG
QLVTVYHDPSSQADSNPNVELPDVRAANYPPIPLHMANISDAKSTNFTSKITICS
SHKQPKTSLNPNALNRPDVCVSANKERINDVPHFVEDPFOPEYMLTKKED
LGSITWEGIVDKCELVYKHSDFVSLDEKIQELISQTFEEDVPLNVAEQVNCVNLG
FNCDCQEMTCVMTNLHQPKWOMLKKHAYVTGIPKSGVYEAEPKALFOEYLVN
SKSRFGKRLTSLIKERFPVVKDLRLILHVAIVVMGVSFFGVKLFPAKNRECEVLHVS
ESYVDGNKIPRTRESYETPNVPTRESYESNNVSVKTESPVSFVKNVRFBMSPS
GRVSGSENMALSLQLDNYIKVEKQGVSONMAELCSKLVTKRMKPIYENHVSIPIL
GHVLFTRKRIAIMPHHFLAALKFKRQEGSGVYVRNPLFSATFFYKLEDMIRKYPF
ESBPPEPESLAESRDLCSFCLDNTINFSVDKLFVSKSDLSYKSDILPLTSTBSNG
QAFARIKIGRAASGIGORDCRVYSGDRLRLVRCRVYVETEGVDCGAGLILARN
ALAGRKIMGIIHAGNTDAGFTPLKEDIDTILSMYPLSSQVANECCQIHPPTCHL
PESTGFVVDLKEKPLVYASGKVISPSPLHGLITTPKTRCOLRDPSPSPVQVLEK
FASCPYVDVABWLENSVAVSNGHLSKSLIENKDLITTSQPSKSPFSEAVSGIDDEEFT
NSVKSSSSGYFPYVDKENSKEKIPKGPBPEDVYTRKAILILROOYEBEISQAKGVR
QGVHFDITDKERKPIYHKAHKTRMFSACGLDLILACKMPLFGVAVSLQKSRNICGISV
GTNVYSIDWTIITANTLKSQPCMIAGDEFGPDSQLODILRAASQVLLNOKSRNICGISV
EBDLVMOQVLESLSSVHLNNNYVMMLKGLPSGFLVLAIIINSIFVLISFNSVQIA
FGVNAKAEPEFVCGIYVAGDHIIVSYEMATNPFOVLEALSFKQILSLSTLEBKD
ATVNA PYSLNVSYLKRRKFLDEDEKROVLAFLSETLLETFMWYKKCVNVLQTTTE
LENSLKELCYHSQVSWDHSIHSEFKCAICLLSSPFLFIDREPARSEVLENV"
<6473 . 9097

QY 61 UGACUUUUUAAGACCCCUUAGUUUAUAGCUUUACCCGCCAGAGUGGGUGGACCGUUCU 120
Db 6346 TACATTTTAAAGACCTTAGTTATTTATTTAGCTTAAACCCCGAGATGGGCTGCAGCGCTTCT 6405
QY 121 GCAUUAUCCAGGGGACCCUAGUGGACCGCUUUGUUUUUGUGGACUUUAGGCUAAAGAU 180
Db 6406 GCAATATCCAGGGGACCTTAGTGCACGCTTGTAGTTTGTGACCTTTAGGCTTAAAGAT 6465
QY 181 UUCACUA 187
Db 6466 TTCACCTA 6472

RESULT 5
LOCUS BD173511 188 bp RNA linear PAT 18-FEB-2003
DEFINITION Novel tertiary structure having ability to accelerate translation activity.
ACCESSION BD173511
VERSION BD173511.1 GI:28414842
KEYWORDS WO 02061080-A/1.
SOURCE Plautia stali intestine virus
ORGANISM Plautia stali intestine virus
VIRUSES; ssRNA positive-strand viruses, no DNA stage;
Dicistroviridae; Crispavirus.
1 (bases 1 to 188)
REFERENCE Nakashima, N. and Kanamori, Y.
AUTHORS Novel tertiary structure having ability to accelerate translation activity.
TITLE Patent: WO 02061080-A 1 08-AUG-2002;
JOURNAL JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF
SERICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE
FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA, YASUSHI KANAMORI
OS Plautia stali intestine virus
PN WO 02061080-A/1
PD 08-AUG-2002
PF 31-JAN-2001 WO 2001JP000641
PI 25-JAN-2001 JP 01P 016746
PC NOBUHIKO NAKASHIMA, YASUSHI KANAMORI
PC C12N15/09, C12N15/86, C12P21/02
CC Novel tertiary structure having ability to accelerate translation activity
FH key Location/Qualifiers
FT source 1.188
/organism='Plautia stali intestine virus'.
location/Qualifiers
1.188
/organism='Plautia stali intestine virus'
/mol_type='genomic RNA'
/db_xref='taxon:64698'

COMMENT
ORIGIN
Query Match 34.4%; Score 64.4; DB 6; Length 188;
Best Local Similarity 39.8%; Pred. No. 8.6e-06;
Matches 74; Conservative 43; Mismatches 66; Indels 3; Gaps 1;
FEATURES
source
location/Qualifiers
1.188
/organism='Plautia stali intestine virus'
/mol_type='genomic RNA'
/db_xref='taxon:64698'

ORIGIN
Query Match 34.4%; Score 64.4; DB 6; Length 188;
Best Local Similarity 39.8%; Pred. No. 8.6e-06;
Matches 74; Conservative 43; Mismatches 66; Indels 3; Gaps 1;
FEATURES
source
location/Qualifiers
1.188
/organism='Plautia stali intestine virus'
/mol_type='genomic RNA'
/db_xref='taxon:64698'

RESULT 6
LOCUS BD177015 188 bp RNA linear PAT 16-APR-2003
DEFINITION Novel translational activity-promoting higher-order structure.
ACCESSION BD177015
VERSION BD177015.1 GI:30014275
KEYWORDS JP 2002306168-A/1.
SOURCE Plautia stali intestine virus
ORGANISM Plautia stali intestine virus
VIRUSES; ssRNA positive-strand viruses, no DNA stage;
Dicistroviridae; Crispavirus.
1 (bases 1 to 188)
REFERENCE Nakashima, N. and Kanamori, Y.
AUTHORS Novel translational activity-promoting higher-order structure
TITLE Patent: JP 2002306168-A 1 22-OCT-2002;
JOURNAL DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND
ENTOMOLOGICAL, HIROSHIMA PREFECTURAL INDUSTRIAL TECHNOLOGY PROMOTION
ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND
FISHERIES
OS Plautia stali intestine virus
PN JP 2002306168-A/1
PD 22-OCT-2002
PF 25-JAN-2001 JP 2001016746
PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI
PC C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02// PC
PC C12N15/09, C12N1/92, C12N15/00, C12N5/00, C12N15/92) CC
Novel translational activity-promoting higher-order structure
FH key Location/Qualifiers
FT source 1.188
/organism='Plautia stali intestine virus'.
location/Qualifiers
1.188
/organism='Plautia stali intestine virus'
/mol_type='genomic RNA'
/db_xref='taxon:64698'

ORIGIN

Query Match 34.4%; Score 64.4; DB 6; Length 188;
Best Local Similarity 39.8%; Pred. No. 8.6e-06;
Matches 74; Conservative 43; Mismatches 66; Indels 3; Gaps 1;
FEATURES
source
location/Qualifiers
1.188
/organism='Plautia stali intestine virus'
/mol_type='genomic RNA'
/db_xref='taxon:64698'

ORIGIN
Query Match 34.4%; Score 64.4; DB 6; Length 188;
Best Local Similarity 39.8%; Pred. No. 8.6e-06;
Matches 74; Conservative 43; Mismatches 66; Indels 3; Gaps 1;
FEATURES
source
location/Qualifiers
1.188
/organism='Plautia stali intestine virus'
/mol_type='genomic RNA'
/db_xref='taxon:64698'

ORIGIN
Query Match 34.4%; Score 64.4; DB 6; Length 188;
Best Local Similarity 39.8%; Pred. No. 8.6e-06;
Matches 74; Conservative 43; Mismatches 66; Indels 3; Gaps 1;
FEATURES
source
location/Qualifiers
1.188
/organism='Plautia stali intestine virus'
/mol_type='genomic RNA'
/db_xref='taxon:64698'

[illegible]

```

source
1. 8797
/organism="Plautia stali intestine virus"
/mol_type="genomic RNA"
/specific host="Plautia stali"
/db_xref="taxon:64698"
571..606
/codon_start=1
/product="nonstructural polypeptide"
/protein_id="BAA21898.1"
/db_xref="GI:2344757"
/translational_table="1"
GKSYASYSIERMIFEDLNHPYGVNVMKRNIOIKRWYVNMVVCUSFDRKVOG
YGVNMPHFLDRHDETSQYNTNHLFRSDVDLVYERNCWFDELSRDVETYSFIL
GLPENIYCNLDQILDRFTSDELFIHLENQFARKPLVFLDSKMFQOHLSEFKK
IDNYSILDTDLKFLVSHMERVSKEFLPGRSTVSLGSDIIVEDPYVVOHQTSX
AYVEPOLNLOSNGNILEDPIRKYTELETFMGQGLGGINITIDIKDQLOAVHQSX
NAYESOMQVKGKYLKFLIKVFLKGLVLSGLKIKDLSALVITFLFLFGKMC
SLDKPIFYSESLSNOCDCDESOFLSLFLDLFLNGCSQMNADPWFVSAOPPL
PSQGHITIVSKDLYVCHERFPRVKLGLPESYSPCVCMIAODIYKRYKNTIL
VLDRLLDLKLFNLYKEGNRFLCTPAFNOQOIYKVINLYNLIDKYSISRGYKNS
LRPPVSLILLGTRGRKTYVPLTVEVATRIYLEBHEGPIITDEIASIYANSEB
EYMGYTGOLITVDEPMQVDSANPNLEFEMIRASHEPIYPLHMANEDKNWME
RSSVTLASNSLTAENLQSKSLNYPALLRLPVLVEVETAPCTGRAPAGOSK
YKFKRYVITLIDESRVSIVASEISYDYLIVLMCLKYDMNTTCOSANTITEMGNR
KQWKAKEBSHLTIBESNNLDAQWLSFVDSYEENDIGYNFDEVTTPQOYIIF
KFEYVAFIDITPDPDEISADKERTTIMWRHESQICEEPIYVLTPLGLVVALG
GYTIFRCFNGETTPLAKESEVILFKPESEKKEVIRCKEISEQKIGSVRSES
EKGAIPHVKLEQSVNATEDSVMOGNVANTLONYSKVMETLVEGSCQONMA
ELISKVCNRYVLAFCVRPGRESRLHILFLKQIKYVPOHPLFSRLKMEESPD
ISLRSIFLRNMETIYICDPLANTNIPENDOGGLNCSVDVETVTKHPIILTVYS
QTEVKSILRSDVCLPFIHVESNKRKYATVATGQSGDSQANYSVAALITELCKM
ROSKRYLQTSAGTCGAPVILIGAKQSGRI CGMHWAGDSQANYSVAALITELCKM
NDLNPITIOSSEMERKMIQNGFDTLPPEPGKFIISAGDEPISASAKTQLRKSVYGEI
APVLTPTMLTPTGLNEMWDPNRYRLSLGRPTLVKMLNLSIKDLRVLIYMEYEL
KSNKESRYPEETACGIDSDPTFNSIKRTSGVGLGCKVKGKQOIFGSDGPFNN
GYNLDRKDVHEIIESLADGISVYHPIIDLDEKALIEKATKELFASPLPYLL
KTLNLSLTLDCMNMORSGEDEGLMAQANSFLAKELSIHEDEVWDQMFLILAAKNR
HGEVEFLSRHAQVAVRETDFFEEBS"
<6193..8628
/notes="Either readthrough of terminal codon of the first
orf for nonstructural polypeptide or use of internal
ribosome binding site. Confirmed N-terminal of capsid
protein precursor starts at 6193 by N-terminal analysis"
/codon_start=1
/product="capsid protein precursor"
/protein_id="BAA22088.1"
/db_xref="GI:2361038"
/translational_table="1"
KIDBUISCAPQTMNESRPASDPREHITVQFLERPVAVATHNGVTEADARNINLIDLEI
PKALLDMNLNKEPDEGFSFSATVAFKQINSQPOAGLLIGALPSKDLGSRNTDVKA
VADKSLIYPIHTLIDIKSEITLSVTPYSPFPQYNVLEBINSNFPKIRKYSVLS
QVODDLVLARAFDIDILGYPTVLPVTPFDLLLOSGETGSPSKAAITIVDAEGS
GKGSYNTIPYKPFVNMQNTVNGRGIOLSLAAGFSGSKQDNQOVNUPPEFNLANL
GVDTGVMSAPAGNALPLMSIGSGSDDELSEFKLLDQOPNIDVESVSTSTPTTLM
STHLCPEFLSKTDIQHROPULIYLSNPFILYMGSLKFTPRPKTYNHSRGLVFS
PFSQTSOSDFNRSAYAVKVMQDLREGBETLSVIPYNTNYSICDMRTTSPRIDATVY
TNPQTIVAHSPCMIAINALPVLQABEELPVSIDCVESGGDFELQAIANGWGVC
FDSASSQLTLQSGDTGSGTGRINVTYNDKVDPSVGNNSLYVDGATLTKMDPKAG
ERWASMRPLKRPSVAFSTGNLTSTYDILRLNLIFFDDPSYSLVDGATLTKMDPKAG
NLISRIYOMVAFYRGNINIKVAPDKGVNPLVYAVISGLTTSNTMSYFSEOVNYS
AKSICERNPYTNSFKRSAYATNQTENVNTQPFNPFLAAGRVASADDPDCGFLGE
PSPSRFPLTKTIS"

```


Db 6006 ACTATGTCATCTTTATTAATAATTAGTTAAATTTGAGAGTTTAAATAATGTTTATATATTC 6065
 Qy 61 UACAUUUUUAAAGCCUUUAGUUUUAGCUUUAACGCCAGAGUGGGUGCAGCGUUCU 120
 Db 6066 TATAGTCTTGAAGGCTGTATATTTATATTAATCAACACAGATGACCGGAGCAGCCCT 6125
 Qy 121 GCAUUAUCCAGGAGCCUAGUGGAGCCUUUUAGUUGAGUCUUUAGGCUAAGAUAU 180
 Db 6126 CCAATATCTATGTATGCTCGTCTCGCTCAAAACATTAAAGTGTGTGCGAAAGAAT 6185
 Qy 181 UUCACU 186
 Db 6186 CTCAC 6191

RESULT 9
 BD173514 188 bp RNA linear PAT 18-FEB-2003
 LOCUS Novel tertiary structure having ability to accelerate translation
 DEFINITION
 ACCESSION BD173514
 VERSION BD173514.1 GI:28414845
 KEYWORDS WO 02061080-A/4.
 SOURCE Cricket paralysis virus
 ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Dicistroviridae; Crispavirus.

REFERENCE 1 (bases 1 to 188)
 AUTHORS Nakashima,N. and Kanamori,Y.
 TITLE Novel tertiary structure having ability to accelerate translation activity
 JOURNAL Patent: WO 02061080-A 4 08-AUG-2002;
 JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF
 SERICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE
 FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA,YASUSHI KANAMORI
 OS Cricket paralysis virus
 PN WO 02061080-A/4
 PD 08-AUG-2002
 PR 25-JAN-2001 JP 01P 016746
 PI NOBUHIKO NAKASHIMA,YASUSHI KANAMORI
 PC C12N15/11,C12N15/86,C12P21/02
 CC Novel tertiary structure having ability to accelerate CC translation activity
 FH Key Location/Qualifiers
 FT source 1..188 /organism='Cricket paralysis virus',
 location/Qualifiers
 1..188 /organism='Cricket paralysis virus'
 /mol_type='genomic RNA'
 /db_xref='taxon:12136'

ORIGIN
 Query Match 24.8%; Score 46.4; DB 6; Length 188;
 Best Local Similarity 36.6%; Pred.No. 0.3;
 Matches 63; Conservative 37; Mismatches 71; Indels 1; Gaps 1;

Qy 15 UGAUUAAGAGUAGAAAUUCCUAGUUUAUUAUUUUUUAUACUGCUACAUUUUUAAGAC 74
 Db 15 TGCCTGTAATAATCAATTTGAGAGGTTAATTAATTAACAAGTAGCTATTTTGTATTTA 74
 Qy 75 CCUUAUUUUUUAUAGCCGACGAGAGUGGGUGCAGGUCUCCUGAUAUUAUCCAGGC 134
 Db 75 GGTTAGCTATTAGCTTTACGTTCCAGAGATCCTTAGTGCGACGCCACAAATATCCAGGA 134
 Qy 135 A-CCUAGUGCAGCCUUGAGUUUUUUAUGAGACUUUAGGCUAAGAUAUUCAC 185
 Db 135 GCCCTCTGCGGTTTTTCAGATTAGTAGTGCAGAAAAACCTAAGAAATTAC 186

RESULT 10

BD177018 188 bp RNA linear PAT 16-APR-2003
 LOCUS Novel translational activity-promoting higher-order structure.
 DEFINITION
 ACCESSION BD177018
 VERSION BD177018.1 GI:30014278
 KEYWORDS JP 2002306168-A/4.
 SOURCE Cricket paralysis virus
 ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Dicistroviridae; Crispavirus.

REFERENCE 1 (bases 1 to 188)
 AUTHORS Nakashima,N. and Kanamori,Y.
 TITLE Novel translational activity-promoting higher-order structure
 JOURNAL Patent: JP 2002306168-A 4 22-OCT-2002;
 DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND
 ENTOMOLOGICAL HIROSHIMA PREFECTURAL INDUSTRIAL TECHNOLOGY PROMOTION
 ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND
 FISHERIES
 OS Cricket paralysis virus
 PN JP 2002306168-A/4
 PD 22-OCT-2002
 PR 25-JAN-2001 JP 2001016746
 PI NOBUHIKO NAKASHIMA,YASUSHI KANAMORI
 PC C12N15/09,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P21/02// PC
 (C12N15/09,C12R1:92),C12N15/00,C12N5/00,C12R1:92) CC
 Novel translational activity-promoting higher-order structure FH
 Key Location/Qualifiers
 FT source 1..188 /organism='Cricket paralysis virus',
 location/Qualifiers
 1..188 /organism='Cricket paralysis virus'
 /mol_type='genomic RNA'
 /db_xref='taxon:12136'

ORIGIN
 Query Match 24.8%; Score 46.4; DB 6; Length 188;
 Best Local Similarity 36.6%; Pred.No. 0.3;
 Matches 63; Conservative 37; Mismatches 71; Indels 1; Gaps 1;

Qy 15 UGAUUAAGAGUAGAAAUUCCUAGUUUAUUAUUUUUUAUACUGCUACAUUUUUAAGAC 74
 Db 15 TGCCTGTAATAATCAATTTGAGAGGTTAATTAATTAACAAGTAGCTATTTTGTATTTA 74
 Qy 75 CCUUAUUUUUUAUAGCCGACGAGAGUGGGUGCAGGUCUCCUGAUAUUAUCCAGGC 134
 Db 75 GGTTAGCTATTAGCTTTACGTTCCAGAGATCCTTAGTGCGACGCCACAAATATCCAGGA 134
 Qy 135 A-CCUAGUGCAGCCUUGAGUUUUUUAUGAGACUUUAGGCUAAGAUAUUCAC 185
 Db 135 GCCCTCTGCGGTTTTTCAGATTAGTAGTGCAGAAAAACCTAAGAAATTAC 186

RESULT 11
 AF218039 9185 bp RNA linear VRL 02-JUL-2000
 LOCUS Cricket paralysis virus nonstructural polyprotein and structural
 DEFINITION polypeptide genes, complete cds.
 ACCESSION AF218039
 VERSION AF218039.1 GI:8895506
 KEYWORDS
 SOURCE Cricket paralysis virus
 ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Dicistroviridae; Crispavirus.

REFERENCE 1 (bases 1 to 9185)
 AUTHORS Wilson,J.E., Powell,M.J., Hoover,S.E. and Sarnow,P.
 TITLE Naturally occurring dicistronic cricket paralysis virus RNA is
 regulated by two internal ribosome entry sites
 JOURNAL Mol. Cell. Biol. 20 (14), 4990-4999 (2000)
 PUBMED 1086656
 REFERENCE 2 (bases 1 to 9185)
 AUTHORS Wilson,J.E., Powell,M.J., Hoover,S.E. and Sarnow,P.

TITLE Direct Submission
JOURNAL Submitted (20-DEC-1999) Microbiology & Immunology, Stanford
University, 299 Campus Drive, Stanford, CA 94305, USA
FEATURES Location/Qualifiers
SOURCE 1. 9185
/organism="Cricket paralysis virus"
/viroion
/mol_type="genomic RNA"
/db_xref="taxon:12136"
709. 6024
/codon_start=1
/product="nonstructural polyprotein"
/protein_id="AAF80998.1"
/db_xref="GI:8895507"
/translation="MSFQQTNNNTNNINSLEIAAOELIAAOFEGNLDGFCFEPYVO
SKPOLDSEECYCHDDPGCCGDKIKKEELRIKLFISDYGTFEEMKLVKQV
NCEHRYGTFEGGLIVSPRFMDLDWQOKLVSNDCECAFLRKQILMSD
VSNPQVQSPVACNDIPRAIRLEKRLQKRDKEI STLKQKQEIKNRIYQGF
DILKAGGEVQGLNGLRICDFLENSLPETLAIQITVLTTRKYNLKEDLKVAL
LVLVRLMWKKYRAALIVILFVMEHYGDKOILDIVDLKKILOTTOAGTELL
EAVYHMPOTCGCLIPAVLAFAIKIKPGQPDNTISRDRPKAIBSGKIVDYC
SEYNLSVDEKAVVGLKELGTQGLDEIHVAKELRHVYLDENKKTLDITAAK
VEDLYKRGKXSEKIPDRIDARITTLPEAKSLYEQVLLSPVGGGPKRPTVWL
TGESIGIKTQMIYPLCIDILEMGIIVKDAVQAQVAYEYEWGQKIVYD
AFQKDDTKENPEIFVIRTCNTEPQHLMAALODKNYSQAEVLVYTNQFOVLE
SITPDAFYNMKTAAVVOIKOKSIWVRNABEENALDVTLLNDEALDLSVAF
OKAFDDSEATKIDIDGEPISYDEPARTICAKKEBKATHQQMLEAASRVVAG
GSESEYDVWDETFYSNLSQGFMAKSLIEMAEPSADSETNAIIEYKNIPEKT
KMSKMTLDEQIALSTKIRLKNKAYKSEHPYTLALGFIWMISAFMYSFPR
TLDDMTTSEVSGSGDNKTOKISKRVVGGSGVTKVTEAVESGSDSKTMK
KITAEVSGSKTKOKORTKVEVKELEKEATQGSDDPAALVLDVLOKTYCL
YEMVWGMKRVLTATPLRGVCMMPHAFITETAKVAPSTNIYFSGPDNDVI
VPSHIAIPAEVEITACTRIHYKDETRDVLYNHRMHCHPRDILKHYKKS
DQGLKREFOGTATFHOSANELCRAVQLOAIRPQOELITHEIDMPESEST
ORDCEYNAPOJQNCGSIYGLYNKMERKLGIMIPGNSVSECYACPLQEGIMG
LNRLEKDPVNIYVQCCFEPSPIDKTMSETEBGFCAIGKSNIVGQVATLTK
SCIYMSKLPITRKPAHLTRRLPNEGELVDPLMKGLKKGQVTLDAIEVSALDVK
OVVLTQYNSMDVNRKRPFLVEATQGTGDDPMKGIARQTSQVYFQWRPLPK
QDMWGEQVDFTSQRAOELRDVEBLINDCAKGIIDQVYVDTLDBERRIEVDK
KTRFSAQPOHVVAFKRYFLPFAVILKNNIDIEIAGNVISTDWERIKRLKKG
NKVIAGDFGNFGLVAGFQSGCKSFYPMKTFNDVNTDGRKMLICGLMTHIV
HSVSYGDNVYMTSHQPSGNPFTVILINCLINSMIRMIWILLAKRAPEQSMKRR
ENVMISYXGDNCLNISDRVVENQITISEOMKIEKHEYEDNKMDYFELSETI
HELKRRFSPSHOIORVAPLOKDIVIEMLMNTRTDITPNEILMMNTAPREIYVHKK
SEYKLSGIEDLMAKGLRPQPOLIFKALINDATMLADEVYVP"
6217. 8904
/codon_start=1
/product="structural polyprotein"
/protein_id="AAF80999.1"
/db_xref="GI:8895508"
/translation="ATPDQKQENSHIENEDKRLMSQEKIYHVSIGITSTALPDI
VNSTNYLDMTRDRHISIDPLSGPIILATNLMSSDPEKOLYANPEVILISA
MYQDLKGVGLRATLVKQVNSQPOQGLMQLYIPYAMNRYLILNETLQGS
GCPITDLELSTGTEVEMRIYVSPHLYNLTGSGFSIYVYVSOHQHVGSTGSI
EYTAHLEDDVQYPTGANIFGNSPYSIARIARTIGERTIEGIMLHKHYKLI
PARVIAAKEIKOLETNNSPSLAIGISELTLSHPVIGNIFSPAMISAAYAL
AKIFGSKPTQSGIKGECKLRGQGRMANPDMDSHMAIESTIEITKGLATSID
EMDLSRVLSIPNDRFTWTSIDVTNLVMDNVSPKRVKYSATITDRFCHEWIV
ANATYWRGSIYVTFKVTQYHGRISIPYVNTTISTGPDVSRQKIVGDR
TSEVSTFVPIASPMYLCIRPSSSMISKNDGALMYNCSIVAVEVQNLVAAQ
NVFEIDVIGVSGPDLFAGPTSPSYPAAGDLTADTKIAETKISQNSNEDR
ITTOCSRIVAGVMDQOIPNEAHOHVPSISIDTHRIINNNWQCNIDIGKYSIO
LIRKFAFGDANTIQADGSSFYVAPYVTSPTKTLSTRNTQPDYIYLYAPRGS
RIKRVATQDGTGTPRKKTNTFTVRMNSLQDSFNSLISSSAIVTVVLPSTJIM
GPTQVQIDPTVEGLIEVEPYNISHITPAVTIDGTPSMEDYKLGKSPCLTFSPR
DSISATNHTIASFMRABGDDFSFMYLLEVPVLVAVARA"

ORIGIN

Query Match 24.8%; Score 46.4; DB 13; Length 9185;
Beet Local Similarity 36.6%; Pred. No. 0.18;
Matches 63; Conservative 37; Mismatches 71; Indels 1; Gaps 1;
QY 15 UGAUUGAAGAAUUCUAGUUUAUUAUUUAUACUGCUAUAUUUUUAAGAC 74

Db 6043 TCGTGGAAATACAAATTTTGAGAGTTTAAATAATACAGAGTATTTTGTATTA 6102
QY 75 CCUAGUUAUUAUCCUUAUACCGCCAGAGUGGGGUGACGCCUCCUGCAUAUCCAGGCG 134
Db 6103 GGTTAGCTATTACCTTACCTTCAGATGCGCCAGGCGCCCAAAATCTCAGAGA 6162
QY 135 A-CCUAGUGAGCGCCUGUGUUAUUGAGCUUUAAGCUAAGAUUUCAC 185
Db 6163 GCCCTCTGCGGGTTTTCAGATTAGGATGCAAAAACCTAAGAAATTTAC 6214

RESULT 12
AC161224/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (10-MAY-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Broad Institute of MIT and Harvard
Center code: WBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: U32200
Center clone name: 32_L_14

* NOTE: This is a 'working draft' sequence. It currently
* consists of 75 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

* be preserved.
1
* 879: 878: contig of 878 bp in length
* 879: gap of unknown length
* 1826: contig of 848 bp in length
* 1827: gap of unknown length
* 2598: contig of 672 bp in length
* 2698: gap of unknown length
* 3205: contig of 507 bp in length
* 3206: gap of unknown length
* 3306: contig of 643 bp in length
* 3948: gap of unknown length
* 4049: contig of 913 bp in length
* 4962: gap of unknown length
* 5062: contig of 870 bp in length
* 5932: gap of unknown length
* 6032: contig of 807 bp in length
* 6839: gap of unknown length
* 6939: contig of 864 bp in length
* 7803: gap of unknown length
* 7903: contig of 726 bp in length
* 8629: gap of unknown length
* 8729: contig of 860 bp in length
* 9589: gap of unknown length
* 9689: contig of 70 bp in length
* 9759: gap of unknown length
* 9859: contig of 1353 bp in length
* 11212: gap of unknown length
* 11312: contig of 965 bp in length
* 12276: gap of unknown length
* 12377: contig of 938 bp in length
* 13315: gap of unknown length
* 13415: contig of 898 bp in length
* 14313: gap of unknown length
* 14413: contig of 1178 bp in length
* 15591: gap of unknown length
* 15691: contig of 897 bp in length
* 15687: gap of unknown length
* 16588: contig of 743 bp in length
* 17431: gap of unknown length
* 17531: contig of 1242 bp in length
* 18773: gap of unknown length
* 18873: contig of 1307 bp in length
* 20180: gap of unknown length
* 20279: contig of 1136 bp in length
* 20280: gap of unknown length
* 21416: contig of 2050 bp in length
* 21516: gap of unknown length
* 23666: contig of 1235 bp in length
* 24901: gap of unknown length
* 25001: contig of 1099 bp in length
* 26100: gap of unknown length
* 26200: contig of 1263 bp in length
* 27463: gap of unknown length
* 27563: contig of 1209 bp in length
* 28771: gap of unknown length
* 28872: contig of 1378 bp in length
* 30249: gap of unknown length
* 30250: contig of 1745 bp in length
* 30350: gap of unknown length
* 32095: contig of 1134 bp in length
* 32195: gap of unknown length
* 33329: contig of 1721 bp in length
* 33429: gap of unknown length
* 35150: contig of 1562 bp in length
* 35250: gap of unknown length
* 36812: contig of 2147 bp in length
* 36912: gap of unknown length
* 39059: contig of 2235 bp in length
* 39159: gap of unknown length
* 41394: contig of 1924 bp in length
* 41494: gap of unknown length
* 43418: contig of 1213 bp in length
* 43518: gap of unknown length
* 44731: gap of unknown length

* 44831: 47142: contig of 2312 bp in length
* 47143: gap of unknown length
* 47243: contig of 1757 bp in length
* 48999: gap of unknown length
* 49000: contig of 2713 bp in length
* 49100: gap of unknown length
* 51813: contig of 2570 bp in length
* 51913: gap of unknown length
* 54483: contig of 2995 bp in length
* 54583: gap of unknown length
* 54883: contig of 1995 bp in length
* 57578: gap of unknown length
* 57678: contig of 1995 bp in length
* 59673: gap of unknown length
* 59773: contig of 1995 bp in length
* 61768: gap of unknown length
* 61868: contig of 2622 bp in length
* 64490: gap of unknown length
* 64589: contig of 1671 bp in length
* 64950: gap of unknown length
* 66261: contig of 1973 bp in length
* 66361: gap of unknown length
* 68334: contig of 1838 bp in length
* 68434: gap of unknown length
* 70271: contig of 1643 bp in length
* 70372: gap of unknown length
* 72015: contig of 1494 bp in length
* 72115: gap of unknown length
* 73609: contig of 1974 bp in length
* 73709: gap of unknown length
* 75683: contig of 2840 bp in length
* 75783: gap of unknown length
* 78623: contig of 2372 bp in length
* 78723: gap of unknown length
* 81095: contig of 2162 bp in length
* 81195: gap of unknown length
* 83357: contig of 2464 bp in length
* 83457: gap of unknown length
* 85920: contig of 2579 bp in length
* 85921: gap of unknown length
* 86021: contig of 2581 bp in length
* 88601: gap of unknown length
* 88699: contig of 1651 bp in length
* 88700: gap of unknown length
* 91281: contig of 2154 bp in length
* 91381: gap of unknown length
* 93032: contig of 3640 bp in length
* 93132: gap of unknown length
* 95285: contig of 3179 bp in length
* 95385: gap of unknown length
* 95386: contig of 1781 bp in length
* 97576: gap of unknown length
* 97676: contig of 1890 bp in length
* 99457: gap of unknown length
* 99557: contig of 3007 bp in length
* 101447: gap of unknown length
* 101448: contig of 3200 bp in length
* 104554: gap of unknown length
* 104654: contig of 4614 bp in length
* 107854: gap of unknown length
* 107954: contig of 2274 bp in length
* 110228: gap of unknown length
* 110328: contig of 4119 bp in length
* 114447: gap of unknown length
* 114546: contig of 4031 bp in length
* 115477: gap of unknown length
* 118578: contig of 4614 bp in length
* 118678: gap of unknown length
* 123292: contig of 3640 bp in length
* 123392: gap of unknown length
* 127032: contig of 3179 bp in length
* 127132: gap of unknown length
* 130311: contig of 4166 bp in length
* 130411: gap of unknown length
* 134576: contig of 4166 bp in length
* 134577: gap of unknown length

Query Match 21.5%; Score 40.2; DB 14; Length 160056;
Best Local Similarity 41.2%; Pred. No. 4.6; Indels 0; Gaps 0;
Matches 35; Conservative 22; Mismatches 28;

QY 37 UAGUUAUUAUUUUUAUACUGCUACUUUUUAAGACCCUUUAUUUAAGCUUUAACCG 96
Db 70733 TGGATTAAATAATCTCAAGGTAAATCTTGAAAGAACTGTCTCATGTAGCTTATCG 70674
QY 97 CCCAGAUUGGGUGUGAGCCGUCUG 121
Db 70673 TCCAGAGATGGGCTGCAGAGTTTATG 70649

RESULT 13
AC131732 212284 bp DNA linear HTG 25-MAY-2004
LOCUS Mus musculus chromosome 15 clone RP23-119H3, WORKING DRAFT
DEFINITION SEQUENCE, 4 unordered pieces.
AC131732
AC131732 GI:47605179
HTG: HTGS_PHASE1, HTGS_DRAFT, HTGS_ACTIVEPIN.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 212284)
AUTHORS Wilson, R.K.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 212284)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 212284)
Wilson, R.K.
Direct Submission
Submitted (25-MAY-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On May 25, 2004 this sequence version replaced gi:22758576.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
Project Information
Center project name: M_BA0119H03
----- Summary Statistics -----
Sequencing vector: M13, 0%
Sequencing vector: plasmid, 100%
Chemistry: Dye-terminator Big Dye, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 210499 bases at least Q40
Consensus quality: 210754 bases at least Q30
Consensus quality: 210991 bases at least Q20

NOTE: This is a 'working draft' sequence. It currently
consists of 4 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 10018: contig of 10018 bp in length
* 10019 10118: gap of unknown length
* 10119 46629: contig of 36511 bp in length
* 46630 46729: gap of unknown length
* 46730 94280: contig of 47551 bp in length
* 94281 94380: gap of unknown length
* 94381 212284: contig of 117904 bp in length.
Location/Qualifiers
1. 212284

/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="15"
/clone="RP23-119H3"
1. 10018
/note="assembly_name:Contig16
clone end:SP6
vector_side:right"
10019. 10118
/estimated_length=unknown
10119. 46629
/note="assembly_name:Contig17"
46630. 46729
/estimated_length=unknown
46730. 94280
/note="assembly_name:Contig18"
94281. 94380
/estimated_length=unknown
94381. 212284
/note="assembly_name:Contig19"

misc_feature
misc_feature
gap
gap
misc_feature
misc_feature
gap
misc_feature

ORIGIN
Query Match 21.5%; Score 40.2; DB 14; Length 212284;
Best Local Similarity 41.2%; Pred. No. 4.4;
Matches 35; Conservative 22; Mismatches 28; Indels 0; Gaps 0;

QY 37 UAGUUAUUAUUUUUAUACUGCUACUUUUUAAGACCCUUUAUUUAAGCUUUAACCG 96
Db 24839 TGGATTAAATAATCTCAAGGTAAATCTTGAAAGAACTGTCTCATGTAGCTTATCG 24898
QY 97 CCCAGAUUGGGUGUGAGCCGUCUG 121
Db 24899 TCCAGAGATGGGCTGCAGAGTTTATG 24923

RESULT 14
AC157478 229783 bp DNA linear ROD 09-AUG-2005
LOCUS Mus musculus BAC clone RP23-33K19 from chromosome 15, complete
DEFINITION Sequence.
AC157478
AC157478.2 GI:71980388
HTG.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 229783)
AUTHORS Zheng, X., Cotton, M., Kozliewicz, A., Haglund, K. and Cordum, H.
TITLE The sequence of Mus musculus BAC clone RP23-33K19
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 229783)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2005) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 229783)
Wilson, R.K.
Direct Submission
Submitted (09-AUG-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Aug 9, 2005 this sequence version replaced gi:59858791.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
Project Information
Center project name: M_BA0033K19
----- Summary Statistics -----

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e. phred quality >=30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence. This finishing standard has slightly changed from the previous Human standard. Specifically, standards for regions of low sequence complexity (such as dinucleotide repeats and small unit tandem repeats) have been relaxed. These regions are very prevalent in the mouse genome, and the return on extended finishing efforts is minimal.

If a sequence meets the criteria of the above statement, it needs no comments or tags. If the criteria are not met, such as ambiguous bases, then the region is duly annotated.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The BAC Library has been constructed by Kazutoyo Osegawa and Minako Tateno in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

This sequence is the entire insert of the clone.

FEATURES

source

```
1..229783
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10090"
  /chromosome="15"
  /clone="RP23-33K19"
  /clone_1fb="RPCI-23"
```

ORIGIN

Query Match 21.5%; Score 40.2; DB 9; Length 229783;

Best Local Similarity 41.2%; Pred. No. 4.3;

Matches 35; Conservative 22; Mismatches 28; Indels 0; Gaps 0;

37 UAGUUAUUAUUUUUAUACUGACAUUUUUUAGACCCUUAUUUUUUUAGUCCG 96

Db 197491 TGGATTTTAAATAATACCTCAAGGTAATCTTGAAGAACTGTGTCATGTGCTTATCG 197432

97 CCCAGAUUGGUGUGUACGCGUUCUG 121

Db 197431 TCCAGATGGCGCTCAGAGTTTATG 197407

RESULT 15

LOCUS

CR848749 164647 bp DNA linear HTG 31-JUL-2005

DEFINITION Danio rerio chromosome 15 clone CH211-147F13, WORKING DRAFT

SEQUENCE 8 unordered pieces.

ACCESSION CR848749 GI:71534791

VERSION CR848749.4 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS Danio rerio (zebrafish)

SOURCE Danio rerio

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 164647)
McLaren,S.
Direct Submission
Submitted (30-JUL-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: http://www.sanger.ac.uk/Projects/D_rerio/fage.shtml#data#light
On Jul 31, 2005 this sequence version replaced gi:54306004.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: ZC147F13
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 160713 bases at least Q40
Consensus quality: 161321 bases at least Q30
Consensus quality: 161959 bases at least Q20
Insert size: 163947; sum-of-contigs
Insert size: 162403; 7.2% error; agarose-fp
Quality coverage: 7.21x in Q20 bases; sum-of-contigs Quality coverage: 7.87x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preerred.

```
1 4964: contig of 4964 bp in length
* 4965 5064: gap of 100 bp
* 5065 26992: contig of 21928 bp in length
* 26993 27092: gap of 100 bp
* 27093 45046: contig of 17954 bp in length
* 45047 45146: gap of 100 bp
* 45147 48001: contig of 2855 bp in length
* 48002 48101: gap of 100 bp
* 48102 52520: contig of 4419 bp in length
* 52521 52620: gap of 100 bp
* 52621 128367: contig of 75747 bp in length
* 128368 128467: gap of 100 bp
* 128468 149059: contig of 20602 bp in length
* 149070 149169: gap of 100 bp
* 149170 164647: contig of 15478 bp in length.
```

FEATURES

source

```
1..164647
  /organism="Danio rerio"
  /mol_type="genomic DNA"
  /db_xref="taxon:7955"
  /chromosome="15"
  /clone="CH211-147F13"
  /clone_1fb="CHORI-211"
  1..4964
  /note="assembly_fragment:00077
  fragment_chain:1
  clone_end:506
  vector_side:left"
  5065..26992
  /note="assembly_fragment:00798
  fragment_chain:1"
  27093..45046
  /note="assembly_fragment:00141
  fragment_chain:1"
  45147..48001
  /note="assembly_fragment:00025"
  48102..52520
  /note="assembly_fragment:00043
  fragment_chain:2"
```

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

This Page Blank (used)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 04:43:02 ; Search time 270.239 Seconds
(without alignments)
4611.840 Million cell updates/sec

Title: US-10-088-750C-2

Perfect score: 187
1 gaataaugugaucugauua.....uaggcuagaagaauucaua 187

Sequence: 1

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21:*

- 1: genebegn1980s:*
- 2: genebegn1980s:*
- 3: genebegn2000s:*
- 4: genebegn2001as:*
- 5: genebegn2001bs:*
- 6: genebegn2002as:*
- 7: genebegn2002bs:*
- 8: genebegn2003as:*
- 9: genebegn2003bs:*
- 10: genebegn2003cs:*
- 11: genebegn2003ds:*
- 12: genebegn2004as:*
- 13: genebegn2004bs:*
- 14: genebegn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	187	100.0	187	8	ABZ80708 Himetobi
2	187	100.0	199	6	AA150545 Himetobi
3	64.4	34.4	188	8	ABZ80707 Plautia s
4	64.4	34.4	197	6	AA150556 CrPV-like
5	64.4	34.4	200	6	AA150544 Plautia s
6	64.4	34.4	430	3	AAZ35832 Plautia s
7	63.8	34.1	281	6	AA150551 CrPV-like
8	46.4	24.8	188	8	ABZ80710 Cricket P
9	46.4	24.8	200	6	AA150547 Cricket P
10	37	19.8	190	6	ABZ80712 Black que
11	37	19.8	202	6	AA150549 Black que
12	36	19.3	189	8	ABZ80709 Drosophila
13	36	19.3	201	6	AA150546 Drosophila
14	34.6	18.5	110000	13	ABD32780_1 Continuation (2 of
15	33.8	18.1	300000	12	ADZ86352 Human prp
16	33.8	18.1	300001	12	ADZ86352 Human prp
17	33.4	17.9	510	13	ADQ54170 Novel can
18	33.4	17.9	190000	10	ADL13752 Osteoarth
19	33.2	17.8	110000	11	ACN44014_1 Continuation (2 of

C	20	33	17.6	861	10	ADF00812 Bacterial
C	21	33	17.6	2655	4	AB13062 Drosophila
C	22	32.8	17.5	2714	13	ADR25713 Breast ca
C	23	32.8	17.5	2714	13	ADP54371 Human pro
C	24	32.6	17.4	349981	10	ADC87619 Human GPC
C	25	32.4	17.3	601	6	ABN63038 Human can
C	26	32.4	17.3	731	4	AA102045 Human rep
C	27	32.4	17.3	2000	11	ACL35363 Rice stre
C	28	32.4	17.3	2792	4	AB119920 Drosophila
C	29	32.4	17.3	3660	4	ABL26664 Drosophila
C	30	32.4	17.3	26493	12	ADJ12386 DNA fragm
C	31	32.4	17.3	46951	10	ADZ13891 Human opt
C	32	32.2	17.2	154	4	AA125563 Probe #15
C	33	32.2	17.2	154	4	ABA71772 Human foe
C	34	32.2	17.2	154	4	AA152107 Probe #20
C	35	32.2	17.2	154	4	ABA37854 Probe #16
C	36	32.2	17.2	154	4	AAK46200 Human bon
C	37	32.2	17.2	154	4	AAK20141 Human bra
C	38	32.2	17.2	154	4	ABZ45919 Human liv
C	39	32.2	17.2	154	6	ABZ20511 Human gen
C	40	32.2	17.2	187	3	AAZ05603 Human sec
C	41	32.2	17.2	450	4	AA116383 Probe #63
C	42	32.2	17.2	450	4	ABA59254 Human foe
C	43	32.2	17.2	450	4	AA139045 Probe #77
C	44	32.2	17.2	450	4	ABA27994 Probe #64
C	45	32.2	17.2	450	4	AAK33250 Human bon

ALIGNMENTS

RESULT 1	ABZ80708	standard; RNA; 187 BP.
ID	ABZ80708	standard; RNA; 187 BP.
AC	ABZ80708;	
DT	15-OCT-2003	(first entry)
DE	Himetobi P virus derived pseudoknot sequence.	
KW	Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;	
KW	albumen; impurity; higher-order structure; intergenic region; IGR-IRIS;	
KW	internal ribosome entry site; se.	
OS	Himetobi P virus.	
XX	MO2003033719-AL.	
XX	24-APR-2003.	
XX	08-OCT-2002; 2002MO-JP010447.	
XX	17-OCT-2001; 2001JP-00319923.	
XX	(NAG-) NAT INST AGROBIOLOGICAL SCI.	
XX	(WAK-) WAKENYAKU KK.	
XX	Nakashima N, Shibuya N, Nishikawa S;	
XX	WPI; 2003-403230/38.	
PT	Cell-free protein synthesis means in wheatgerm system to establish	
PT	overexpression of target gene with base sequence sustaining translation	
PT	activity and function promotion, for producing useful proteins.	
PS	Claim 1; Page 32; 39pp; Japanese.	
CC	The invention relates to a cell-free protein synthesis system derived	
CC	from wheatgerm where there is substantial exclusion of wheatgerm embryo	
CC	albumen impurities. The novel system uses a sequence having a higher-	
CC	order RNA structure that promotes translation activity. The higher-order	
CC	sequence is preferably a "pseudoknot", especially derived from a range of	

CC virusas(AB280707-AB280713). This sequence represents the "pseudocinct"
CC higher-order sequence from the himecobi P virus. The sequence is used in
CC a construct which may also include an intergenic region and internal
CC ribosome entry site (IGR-IRSE). The method is applicable in producing
CC useful proteins
XX
XX Sequence 187 BP; 53 A; 32 C; 39 G; 0 T; 63 U; 0 Other;

Query Match	100.0%;	Score 187;	DB 8;	Length 187;
Best Local Similarity	100.0%;	Pred. No. 4.7e-45;		
Matches 187;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 1 GAAAAUGUGUGAUCUGAUAUAGAGUAGAAAUUCCUAGUAUAUAUUAUUUAUACUGC 60

Db 1 GAAAATGUGUGAUCUGAUAUAGAGUAGAAAUUCCUAGUAUAUAUUAUUUAUACUGC 60

Qy	61	UACAUUUUUUAAGACCCUUAAGUUAUUUAGCUUUAACCGCCAGAGUUGGGGUGACAGCUCU	120
Db	61	UACAUUUUUUAAGACCCUUAAGUUAUUUAGCUUUAACCGCCAGAGUUGGGGUGACAGCUCU	120

Dy 121 GCAAUUAUCCAGGGGACCTUAGGUGCAGCCTTUGAUGUUUAGUGGAACTTTUAGGCUAAAAGAAT 180

Dz 121 GCAAUUAUCCAGGGGACCTUAGGUGCAGCCTTUGAUGUUUAGUGGAACTTTUAGGCUAAAAGAAT 180

Qy	181	UUCACUA	187
Db	181	UUCACUA	187

RESULT 2
AAL50545
ID AAL50545 standard; RNA; 199 BP.

AC	AAL50545;
XX	
DT	19-DEC-2002 (First entry)

DE	Himetobi P virus RNA sequence.
XX	
KW	CrPV-like virus; ss; higher-order

KW	protein synthesis; structural analysis.
XX	
OS	Himetrobi P virus.

PN	WO200261080-A1.
XX	
PD	08-AUG-2002.

PF	31-JAN-2001; 2001WO-JP000641.
XX	
PR	25-JAN-2001; 2001JP-00016746.

PA (NAAG-) NAT INST AGROBIOLOGICAL SCI.
XX
PI Nakashima N, Kanamori Y;

DR	WPI; 2002-627482/67.
XX	
PT	Translational activi

Translational activity-promoting higher-order structure of CrpY-like
viruses for protein translation when suitably initiated, useful in
synthesis of proteins and polypeptides of foreign species for application
in drugs.

Claim 1, Fig 1-2, 3bpp, Japanese.

Claim 1; Fig 1-2; 38pp; Japanese.

The invention comprises seven RNA sequences (CrpV-like viuses) which have a higher-order structure that sustains translational activity-promoting function. The RNA sequences of the invention are useful in the synthesis of proteins and polypeptides for application in developing and producing drugs. The RNA sequences of the invention are also useful in basic research of protein synthesis and structural analysis by the gene recombinant technique. The present nucleotide represents a Himechobi P

CC XX SQ	virus RNA sequence of the invention
Sequence 199 BP; 60 A; 33 C; 40 G;	

Query Match	100.0%;	Score 187;	DB 6;	Length 199;
Best Local Similarity	100.0%;	Pred. No. 4.8e-45;		
Matches 187; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy 1 GAAAATGUGUGAUCUCGATUAGAGUAA GAAAAUUCUGAUGUAUAUUTUUAAATACGCG 60

Db 1 GAAAATGUGUGAUCUCGATUAGAGUAA GAAAAUUCUGAUGUAUAUUTUUAAATACGCG 60

QY 61 UACUUUUUUUAAGACCCUUAAGUUUUUUAGCUUUAACCCGCCAGGAUUGGSGUGACGCGUUCU 120

Db 61 UACAUUUUUUAAGACCCUUAAGUUUUUUAGCUUUAACCCGCCAGGAUUGGSGUGACGCGUUCU 120

Oy 121 GCAUUAUCCAGGGGCACCUAGGUGCAAGCCUUTUUAUGUGGACUUTUUAAGCCUAAAGAU 180
 |||||
 Db 121 GCAUUAUCCAGGGGCACCUAGGUGCAAGCCUUTUUAUGUGGACUUTUUAAGCCUAAAGAU 180

Qy	181	UUCACUA	187
Db	181	UUCACUA	187

RESULT 3
ABZ80707
ID ABZ80707 standard; RNA; 188 BP.

AC	ABZ80707;
XX	
DT	15-OCT-2003 (first entry)

DE Plautia stali intestine virus derived pseudoknot sequence.
XX
KW Pseudoknot; secondary structure; cell-free protein synthesis

KW internal ribosome entry site; ss.
XX
OS Plautia stali intestine virus.

PN WO2003033719-A1.
XX
PD 24-APR-2003.

PF	08-OCT-2002; 2002WO-JP010447.
XX	
PR	17-OCT-2001; 2001JP-00319923.

PA (NAAG-) NAT INST AGROBIOLOGICAL SCI.
PA (WAKE-) WAKENYAKU KK.
XX

XX WPI; 2003-403230/38.
DR
XX

PT Cell-free protein synthesis means in wheatgerm system to establish
PT overexpression of target gene with base sequence sustaining translation
PT activity and function promotion, for producing useful proteins.
XX
PS Claim 1, Page 31-32, 39pp, Japanese.

Claim 1; Page 31-32; 39pp; Japanese.

The invention relates to a cell-free protein-synthesis system derived from wheatgerm where there is substantial exclusion of wheatgerm embryo abumen impurities. The novel system uses a sequence having a higher-order RNA structure that promotes translation activity. The higher-order sequence is preferably a "pseudocistron", especially derived from a range of viruses (AB280707-AB380713). This sequence represents the "pseudocistron" higher-order sequence from the plautia stail intestine virus. The sequence is used in a construct which may also include an intergenic region and internal ribosome entry site (IcR-ires). The method is applicable in producing useful proteins

applicable in producing useful proteins

SQ	Sequence	188 BP;	57 A;	33 C;	35 G;	0 T;	63 U;	0 Other;
Query Match		34.4%	Score 64.4	DB 8;	Length 188;			
Best Local Similarity		62.9%	Pred. No.	5.3e-09;				
Matches 117;	Conservative	0;	Mismatches 66;	Indels 3;	Gaps 1			
QY	AAUGUGUAGUCGUAUUAGAAGUAG---AAAUUCCUGAUGUAAUAUUUUUUUAUAUACUCC	60						
Db	ACUAAUGUAGUCUUUAUAAAUAUVAGGCUAAAUUUUCGAGGUAAAAUAUGUUUAUAUUUC	61						
OY	UACAUUUUUAAGACCUGUAAGUUUAUAGCUNUACCGCCAGAGUUGGGUGUGCAAGCUUCCU	120						
Db	UAUAUCUUAGAAGGUCUUGUAUAUUUAUAUACUAAACACACAAAGUAGAACCGAGACAGCCU	121						
OY	GCAAUAUCCAGGCGCACUAGUGUGACGCCUUGUAUGUUUAUGGAUCUUUAGGCUAAAAGAU	180						
Db	CCAUAUUCUAAUGUGUAGCCUUGUGUCUGCCUCAAACAUUAAGUGUGUUGUGCGMAAAAGAU	181						
OY	181 UUCACU 186							
Db	182 CUCACU 187							
RESULT 4								
AAL50556								
ID	AAL50556 standard; RNA; 197 BP.							
XX	AAL50556;							
DT	19-DEC-2002 (first entry)							
XX	CrPV-like virus-related RNA sequence #6.							
DE								
XX	CrPV-like virus; ss; higher-order structure; drug development;							
KM	drug production; translational activity-promoting function;							
KW	protein synthesis; structural analysis.							
XX	Unidentified.							
OS								
XX								
FH	Key							
FT	misc_binding							
FT	location/Qualifiers							
FT	1..7							
FT	/tag= a							
FT	/bound_molecly= "CrPV-like virus-related RNA sequence #6"							
FT	/note= "Forms a double-stranded region with bases 68-62 of itself"							
FT	14..57							
FT	/tag= b							
FT	35..39							
FT	/tag= c							
FT	/bound_molecly= "CrPV-like virus-related RNA sequence #6"							
FT	/note= "Forms a double-stranded region with bases 142-138 of itself"							
FT	62..68							
FT	/tag= d							
FT	/bound_molecly= "CrPV-like virus-related RNA sequence #6"							
FT	/note= "Forms a double-stranded region with bases 7-1 of itself"							
FT	74..77							
FT	/tag= e							
FT	/bound_molecly= "CrPV-like virus-related RNA sequence #6"							
FT	/note= "Forms a double-stranded region with bases 110-107 of itself"							
FT	80..91							
FT	/tag= f							
FT	97..135							
FT	/tag= g							
FT	107..110							
FT	/tag= h							
FT	/bound_molecly= "CrPV-like virus-related RNA sequence #6"							
FT	/note= "Forms a double-stranded region with bases 77-74 of itself"							
FT	138..142							
FT	/tag= i							
FT	misc_binding							

[illegible]

QY 181 UUCACU 186
Db 182 CUCACU 187

RESULT 5
ID AAL50544 standard; RNA; 200 BP.
XN AAL50544;
XX AAL50544;
XX 19-DEC-2002 (first entry)
XX
XX Plautia stali intestine virus RNA sequence.
DB
XX CrPV-like virus; ss; higher-order structure; drug development;
XX drug production; translational activity-promoting function;
XX protein synthesis; structural analysis.
XX
XX Plautia stali intestine virus.
XX
XX WO200261080-A1.
XX
XX 08-AUG-2002.
XX
XX 31-JAN-2001; 2001WO-JP000641.
XX
XX 25-JAN-2001; 2001JP-00016746.
XX
XX (NAG-) NAT INST AGROBIOLOGICAL SCI.
XX
XX Nakashima N, Kanamori Y;
XX
XX WPI; 2002-627482/67.
XX
XX Translational activity-promoting higher-order structure of CrPV-like
XX viruses for protein translation when suitably initiated, useful in
XX synthesis of proteins and polypeptides of foreign species for application
XX in drugs.
XX
XX Claim 1; Fig 1-2; 38pp; Japanese.
XX
XX The invention comprises seven RNA sequences (CrPV-like viruses) which
XX have a higher-order structure that sustains translational activity-
XX promoting function. The RNA sequences of the invention are useful in the
XX synthesis of proteins and polypeptides for application in developing and
XX producing drugs. The RNA sequences of the invention are also useful in
XX basic research of protein synthesis and structural analysis by the gene
XX recombinant technique. The present nucleotide represents a Plautia stali
XX intestine virus RNA sequence of the invention
XX

Sequence 200 BP; 66 A; 34 C; 37 G; 0 T; 63 U; 0 Other;

Query Match 34.4%; Score 64.4; DB 6; Length 200;
Best Local Similarity 62.9%; Pred. No. 5.4e-09;
Matches 117; Conservative 0; Mismatches 66; Indels 3; Gaps 1;

QY 4 AAUGUGAUCUAGUUAAGUAG--AAAUUCCUAGUUAUAUUUUUUUAUACUCG 60
Db 2 ACUAGUGAUCUUAUUAAAUAUAGUUAAAUAUUUGAGUUAAAUAUUUUUAUADUGC 61

QY 61 UACAUUUUUUUAAGACCUUAAGUUUAUACUUAACCGCCAGAGUGGAGUGCAGGUCUCU 120
Db 62 UAUUGUCUUAAGAGUCUUGUAUUUAUUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUA 121

QY 121 GCAUUAUCCAGGAGCAGCUAGUGGAGCCUUGUAGUUUUUAUGAUAUAUAUAUAUAUAUA 180
Db 122 CCAUUAUCCUAGUUAAGCUCGUCGCUCAAAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUA 181

QY 181 UUCACU 186
Db 182 CUCACU 187

RESULT 6
ID AA235832 standard; DNA; 430 BP.
XX AA235832;
XX
XX AA235832;
XX
XX 03-FEB-2000 (first entry)
XX
XX Plautia stali intestine virus translation promoting sequence.
XX
XX Plautia stali intestine virus; PSIV; translation; promoter; synthesis;
XX ss.
XX
XX Plautia stali intestine virus.
XX
XX JP11290084-A.
XX
XX 26-OCT-1999.
XX
XX 10-APR-1998; 98JP-00114428.
XX
XX 10-APR-1998; 98JP-00114428.
XX
XX (NOR) NORINSUISANSO SANSHI KONCHU.
XX
XX WPI; 2000-016983/02.
XX
XX A DNA promoting translation activity to a protein - and synthesis of a
XX protein efficiently from a protein gene by using the DNA.
XX
XX
XX Claim 4; Page 2; 5pp; Japanese.
XX
XX The present sequence represents a Plautia stali intestine virus (PSIV)
XX translation promoting sequence. Also described in the present invention
XX are: (a) a plasmid having a gene for synthesizing the objective protein
XX downstream of the above DNA base sequence; (b) a plasmid in which the
XX above DNA base sequence is inserted between the promoter sequence and the
XX translation starting point of the objective gene; (c) a transformant in
XX which the above plasmid is introduced to a host cell, a recombinant
XX baculovirus having a gene for synthesizing the objective protein
XX downstream of the above DNA base sequence; (d) a recombinant baculovirus
XX in which the above DNA base sequence is inserted between the promoter
XX sequence and the translation starting point of the objective gene; (e) a
XX method for synthesizing efficiently a protein from a protein gene
XX positioned downstream by taking a DNA corresponding to a base sequence
XX promoting translation activity in the gene base sequence of an insect RNA
XX virus into a plasmid. The translation activating DNA is used for
XX synthesizing the encoded protein
XX

Sequence 430 BP; 131 A; 80 C; 93 G; 126 T; 0 U; 0 Other;

Query Match 34.4%; Score 64.4; DB 3; Length 430;
Best Local Similarity 39.8%; Pred. No. 6.7e-09;
Matches 74; Conservative 43; Mismatches 66; Indels 3; Gaps 1;

QY 4 AAUGUGAUCUAGUUAAGUAG--AAAUUCCUAGUUAUAUUUUUUUAUACUCG 60
Db 236 ACATATGATCTTATTAATTAATTTGAGGTTAAATTAATTAATTAATTAATTAATTAATTC 295

QY 61 UACAUUUUUUUAAGACCUUAAGUUUAUACUUAACCGCCAGAGUGGAGUGCAGGUCUCU 120
Db 296 TATATGCTTGAAGGTCTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 355

QY 121 GCAUUAUCCAGGAGCAGCUAGUGGAGCCUUGUAGUUUUUAUGAUAUAUAUAUAUAUAUA 180
Db 356 CCAATATCTGATGATACCTGCTGCTCAAAATTAATTAATTAATTAATTAATTAATTAATTA 415

QY 181 UUCACU 186
Db 416 CTCACU 421

RESULT 7
 ID AAL50551 standard; RNA; 281 BP.
 AC AAL50551,
 XX
 DT 19-DEC-2002 (first entry)
 XX
 DE CrPV-like virus-related RNA sequence #1.
 XX
 KM CrPV-like virus; ss; higher-order structure; drug development;
 KM drug production; translational activity-promoting function;
 KM protein synthesis; structural analysis.
 XX
 OS Unidentified.
 XX
 FH Key
 FT stem_loop Location/Qualifiers
 FT 1..34
 FT /*tag= a
 FT stem_loop 37..71
 FT /*tag= b
 FT stem_loop 81..148
 FT /*tag= c
 FT stem_loop 154..176
 FT /*tag= d
 FT stem_loop 177..215
 FT /*tag= e
 FT stem_loop 226..254
 FT /*tag= f
 FT stem_loop 258..265
 FT /*tag= g
 XX
 PN W0200261080-A1.
 XX
 PD 08-AUG-2002.
 XX
 PF 31-JAN-2001; 2001WO-JP000641.
 XX
 PR 25-JAN-2001; 2001JP-00016746.
 XX
 PA (NAG-) NAT INST AGROBIOLOGICAL SCI.
 XX
 PI Nakashima N, Kanamori Y;
 XX
 DR WPI; 2002-627482/67.
 XX
 PT Translational activity-promoting higher-order structure of CrPV-like
 PT viruses for protein translation when autably initiated, useful in
 PT synthesis of proteins and polypeptides of foreign species for application
 PT in drugs.
 XX
 PS Example 1; Fig 3; 38pp; Japanese.
 XX
 CC The invention comprises seven RNA sequences (CrPV-like viruses) which
 CC have a higher-order structure that sustains translational activity-
 CC promoting function. The RNA sequences of the invention are useful in the
 CC synthesis of proteins and polypeptides for application in developing and
 CC producing drugs. The RNA sequences of the invention are also useful in
 CC basic research of protein synthesis and structural analysis by the gene
 CC recombinant technique. The present nucleotide represents an RNA sequence
 CC that was used in an example of the invention
 CC
 SQ Sequence 281 BP; 89 A; 52 C; 57 G; 0 T; 83 U; 0 Other;
 XX
 Query Match 34.1%; Score 63.8; DB 6; Length 281;
 Beat Local Similarity 64.6%; Pred. No. 8.9e-09;
 Matches 95; Conservative 0; Mismatches 52; Indels 0; Gaps 0
 40 UUAUUAUUUUUUUAUACUGCUACAUUUUUUAAGACCCUUAUGUAUUUUUACUUAACCGCC 99
 121 UAAUAAAGUUUUUAUU 180

Oy	100	AGAGUUGGAGUGACGCGUUCUGCAUAUUCAGGCGACCUAAGUGCAGCGUUGUUA	159
Db	181	AAGAUUGACCGAGACGACCCUCAAUAUCUGUACCCUGUCUCUCAAACAUAUA	240
Oy	160	GUGGACCUUUGCGCUAAAGAAUUUUCACU	186
Db	241	GUGUGUGUGGCGGAAGAUAUUCUACU	267
RESULT 8			
ABZ80710	ID	ABZ80710 standard; RNA; 188 BP.	
XX	AC	ABZ80710;	
XX	DT	15-OCT-2003 (first entry)	
XX	DE	Cricket paralysis virus derived pseudoknot sequence.	
XX	KM	Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;	
KM	KM	albumen; impurity; higher-order structure; intergenic region; IGR-IRES;	
XX	XX	internal ribosome entry site; ss.	
OS	OS	Cricket paralysis virus.	
XX	PN	WO2003033719-A1.	
XX	PD	24-APR-2003.	
XX	PF	08-OCT-2002; 2002WO-JP010447.	
XX	PR	17-OCT-2001; 2001JP-00319923.	
XX	PA	(NAG-) NAT INST AGRONOMIC SCI.	
XX	PA	(WAKE-) WAKENYAKU KK.	
XX	PI	Nakashima N, Shibuya N, Nishikawa S;	
XX	PI	WPI; 2003-403230/38.	
XX	PT	Cell-free protein synthesis means in wheatgerm system to establish	
PT	PT	overexpression of target gene with base sequence sustaining translation	
XX	PS	activity and function promotion, for producing useful proteins.	
XX	PS	Claim 1; Page 33; 39pp; Japanese.	
XX	XX	The invention relates to a cell-free protein synthesis system derived	
CC	CC	from wheatgerm where there is substantial exclusion of wheatgerm embryo	
CC	CC	albumen impurities. The novel system uses a sequence having a higher-	
CC	CC	order RNA structure that promotes translation activity. The higher-order	
CC	CC	sequence is preferably a "pseudoknot", especially derived from a range of	
CC	CC	viruses (ABZ80707-ABZ80713). This sequence represents the "pseudoknot"	
CC	CC	higher-order sequence from the Cricket paralysis virus. The sequence is	
CC	CC	used in a construct which may also include an intergenic region and	
CC	CC	internal ribosome entry site (IGR-IRES). The method is applicable in	
CC	CC	producing useful proteins	
SO	SO	Sequence 188 BP; 56 A; 33 C; 36 G; 0 T; 63 U; 0 Other;	
Oy	Query Match	24.8%; Score 46.4; DB 8; Length 188;	
Db	Best Local Similarity	58.1%; Pred. No. 0.001;	
	Matches 100; Conservative	0; Mismatches 71; Indels 1; Gaps 1;	
Oy	15	UGAUUAGAAGUAGAAAAAUUCCUAAGUUAUAUUAUUUAUACUGUACAUUUUUAAGAC	74
Db	15	UGCUUGUAAAACAUAUUUGAGAGUUAUAUAUACAGUAGUCUAUUUUUGUAUUUA	74
Oy	75	CCUUAAGUUUAUUAACUUAACCGCCAGAUUGGCGUGCAGCGUUCUGCAAUAUCCAGGCG	134
Db	75	GGUUAGCUUAUUAGCUUUAAGUUCUACAAGAUUCUUAUGGACGCCCAUAUUCACAGAA	134
Oy	135	A-CCUAGUGCAGCCUUGUGUAUUUUUGAGACUUUAGCUUAAGAAUUUUCAC	185

DT 07-AUG-2003 (revised)
DT 19-DEC-2002 (first entry)
XX
XX Black queen-cell virus RNA sequence.
XX
XX CrPV-like virus; ss, higher-order structure; drug development;
KM drug production; translational activity-promoting function;
KM protein synthesis; structural analysis.
XX
XX Black queen cell virus.
OS
XX WO200261080-A1.
XX
XX 08-AUG-2002.
XX
XX 31-JAN-2001; 2001WO-JP000641.
XX
XX 25-JAN-2001; 2001JP-00016746.
XX
XX (NAG-) NAT INST AGROBIOLOGICAL SCI.
XX
XX Nakashima N, Kanamori Y;
PI
DR WPI; 2002-627482/67.
XX
XX Translational activity-promoting higher-order structure of CrPV-like
PT viruses for protein translation when suitably initiated, useful in
PT synthesis of proteins and polypeptides of foreign species for application
PT in drugs.
XX
XX Claim 1; Fig 1-2; 38pp; Japanese.
XX
XX The invention comprises seven RNA sequences (CrPV-like viruses) which
CC have a higher-order structure that sustains translational activity-
CC promoting function. The RNA sequences of the invention are useful in the
CC synthesis of proteins and polypeptides for application in developing and
CC producing drugs. The RNA sequences of the invention are also useful in
CC basic research of protein synthesis and structural analysis by the gene
CC recombinant technique. The present nucleotide represents a Black queen-
CC cell virus RNA sequence of the invention. (Updated on 07-AUG-2003 to
CC correct OS field.)
XX
XX Sequence 202 BP; 63 A; 38 C; 42 G; 0 T; 59 U; 0 Other;
SQ
Query Match 19.8%; Score 37; DB 6; Length 202;
Best Local Similarity 53.5%; Pred. No. 0.62;
Matches 100; Conservative 0; Mismatches 85; Indels 2; Gaps 1;
QY 2 AAAAUGUGUGAUCUGAUAUAGAAUAAUUCUGUGUAUAUUAUUUUUAUAUCUGU 61
DB 4 ACAUUGUGAUCUGUGCGAGGCAAAUUGGCAAGUAUAAUUCGCAAGUAGUGCU 63
QY 62 ACAUUUUUAAGACCCUUAUUAUUAAGCUUUAACGCCCGAGAUU--GGGUGCCAGCGUCC 119
DB 64 AUUGUGGAUUAACCGUAACUAUUAUAGGUUAAGGUUUAUUGGUAAGGAGAGCCCU 123
QY 120 UGCAUAUUAACGAGGAGCAGGUGGAGCGGUGUAUUAUUAUUAUUAUUAUUAUUA 179
DB 124 AUCAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 183
QY 180 UUUUACU 186
DB 184 UUUUACU 190
XX
XX RESULT 12
ABZ80709 ID ABZ80709 standard; RNA; 189 BP.
XX
XX ABZ80709;
XX
XX 15-OCT-2003 (first entry)
XX

DE Drosophila C virus derived pseudoknot sequence.
XX
XX Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;
KM albumen; impurity; higher-order structure; intergenic region; IGR-IRES;
KM internal ribosome entry site; ss.
XX
XX Drosophila C virus.
OS
XX WO2003033719-A1.
XX
XX 24-APR-2003.
XX
XX 08-OCT-2002; 2002WO-JP010447.
XX
XX 17-OCT-2001; 2001JP-00319923.
XX
XX (NAG-) NAT INST AGROBIOLOGICAL SCI.
XX
XX (WAKE-) WAKENYAKU KK.
XX
XX Nakashima N, Shibuya N, Nishikawa S;
PI
DR WPI; 2003-403230/38.
XX
XX Cell-free protein synthesis means in wheatgerm system to establish
PT overexpression of target gene with base sequence sustaining translation
PT activity and function promotion, for producing useful proteins.
XX
XX Claim 1; Page 32; 39pp; Japanese.
XX
XX The invention relates to a cell-free protein synthesis system derived
CC from wheatgerm where there is substantial exclusion of wheatgerm embryo/
CC albumen impurities. The novel system uses a sequence having a higher-
CC order RNA structure that promotes translation activity. The higher-order
CC sequence is preferably a "pseudoknot", especially derived from a range of
CC viruses (ABZ80707-ABZ80713). This sequence represents the "pseudoknot"
CC higher-order sequence from the Drosophila C virus. The sequence is used
CC in a construct which may also include an intergenic region and internal
CC ribosome entry site (IGR-IRES). The method is applicable in producing
CC useful proteins
XX
XX Sequence 189 BP; 57 A; 30 C; 34 G; 0 T; 68 U; 0 Other;
SQ
Query Match 19.3%; Score 36; DB 8; Length 189;
Best Local Similarity 58.3%; Pred. No. 1.2;
Matches 63; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
QY 53 AAUAUCUGAUAUUAUUAAGACCCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 112
DB 54 AGUAGUGCUAUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 113
QY 113 GCGUUCUGCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 160
DB 114 GCACCCCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 161
XX
XX RESULT 13
AAL50546 ID AAL50546 standard; RNA; 201 BP.
XX
XX AAL50546;
XX
XX 19-DEC-2002 (first entry)
XX
XX Drosophila C virus RNA sequence.
XX
XX CrPV-like virus; ss, higher-order structure; drug development;
KM drug production; translational activity-promoting function;
KM protein synthesis; structural analysis.
XX
XX Drosophila C virus.
OS
XX WO200261080-A1.
XX

, Fri Dec 23 12:17:45 2005

us-10-088-750c-2.rng

Page 9

Search completed: December 22, 2005, 08:40:28
Job time : 273.239 secs

This Page Blank (uspo®)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 07:31:33 ; Search time 1956.83 Seconds
(without alignments)
4471.109 Million cell updates/sec

Title: US-10-088-750C-2

Perfect score: 187
Sequence: 1 gaagaugugaugacugauua.....uaggcuagaagaauucacua 187

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: gb_esc1.*
2: gb_esc2.*
3: gb_esc3.*
4: gb_hlc.*
5: gb_esc4.*
6: gb_esc5.*
7: gb_esc6.*
8: gb_esc7.*
9: gb_g881.*
10: gb_g882.*
11: gb_g883.*

Pred. No. is the number of results predicted to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37.8	20.2	261	1	AV333583 AV333583
2	37.6	20.1	920	10	CL076817 CH216-140
3	37	19.8	622	10	CE747951 tigr-g88-CL641258 CH213-8F1
4	37	19.8	1106	10	CL641258 CH213-8F1
5	36.8	19.7	567	9	B2349373 hg90h10.9
6	36.8	19.7	778	9	CC934069 ZMMBRC054
7	36.8	19.7	878	10	CL043477 CH216-57L
8	36.6	19.6	420	7	CN522916 G00123.B3
9	36.4	19.5	378	10	CL279525 Ggal_93C
10	36	19.3	433	7	CO153155 EN01401.5
11	36	19.3	486	7	CO153454 EN02010.5
12	36	19.3	490	7	CO153765 EN02756.5
13	36	19.3	539	7	CO337794 EN15317.5
14	35.8	19.1	638	9	B2344795 hg90h10.b
15	35.8	19.1	666	9	BH876088 hr28h10.b
16	35.8	19.1	794	9	CC719017 OGCWC77TH
17	35.8	19.1	917	10	CG937324 MBMRC39TR
18	35.6	19.0	800	3	BQ118465 EST604041
19	35.4	18.9	678	9	AZ210959 SP_0154_B
20	35.4	18.9	767	11	AZ210046 SP_0153_A
21	35.4	18.9	1041	9	AZ210046 SP_0153_A
22	35.4	18.9	1101	10	CNS00DIF AL070663 Drosophila

C 23	35	18.7	500	9	AZ174007	AZ174007 SP_0125_B
C 24	35	18.7	602	2	BG079884	BG079884 H3046C05-
C 25	35	18.7	636	10	BX233811	BX233811 Danio rer
C 26	35	18.7	781	10	BX197133	BX197133 Danio rer
C 27	35	18.7	846	9	BZ394884	BZ394884 EINH196TR
C 28	35	18.7	935	9	AQ743050	AQ743050 HS_5387_B
C 29	34.4	18.4	451	8	CKX007796	CKX007796 I030d09.b
C 30	34.4	18.4	467	3	BP032244	BP032244 BP032244
C 31	34.4	18.4	557	1	AV767508	AV767508 AV767508
C 32	34.4	18.4	593	9	AZ632412	AZ632412 IM0487H04
C 33	34.4	18.4	602	7	CJ388738	CJ388738 CJ388738
C 34	34.4	18.4	616	11	DR1L9S	DR1L9S
C 35	34.4	18.4	763	10	BX213248	BX213248 Danio rer
C 36	34.4	18.4	793	10	BX140865	BX140865 Danio rer
C 37	34.4	18.4	810	10	BX16081	BX16081 Danio rer
C 38	34.2	18.3	309	10	CE597758	CE597758 tigr-g88-
C 39	34.2	18.3	383	9	AQ028987	AQ028987 CTR-HSP-2
C 40	34.2	18.3	509	7	CN741858	CN741858 SAL_US008
C 41	34.2	18.3	518	3	BM983214	BM983214 UI-CF-BN1
C 42	34.2	18.3	525	7	CN655491	CN655491 SAL_US005
C 43	34.2	18.3	525	7	CN747758	CN747758 SAL_US008
C 44	34.2	18.3	526	7	CN655490	CN655490 SAL_US005
C 45	34.2	18.3	696	3	BU444519	BU444519 BU444519

ALIGNMENTS

RESULT 1
AV333583
LOCUS
DEFINITION
AV333583 RIKEN full-length enriched, adult male medulla oblongata
MUS musculus cDNA clone 6330548020 3', mRNA sequence.

ACCESSION
AV333583
VERSION
KEYWORDS
SOURCE
ORGANISM
MUS musculus (house mouse)

REFERENCE

AUTHORS

Komuro, H., Alizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Iehli, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, U., Kikuchi, N., Kojima, Y., Koya, S., Kuwabara, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takehashi, F., Tateo, M., Tomimaga, N., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Yamamoto, M., and Hayashizaki, Y.
RIKEN Mouse ESTs (Komuro, H., et al. 1999)
Unpublished (1999)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/
Sasaki, N., Izawa, M., Watabiki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.

Oy 122 CAUAUCC 129
 Db 411 CATCATGC 404

RESULT 6
 CC934069/c
 LOCUS
 DEFINITION ZM5BHC0547L03r ZM5BHC Zea mays genomic clone ZM5BHC0547L03 3',
 CC934069
 genomic survey sequence.
 ACCESSION CC934069
 VERSION CC934069.1 GI:33585167
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 778)
 Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
 Rouzard,K., Fuks,G., Yu,Y., Wing,R. and Messing,J.
 Sequencing of the maize genome at PCR (2003D)
 Unpublished (2003)
 Contact: Bharti,A.K.
 Dr.Joachim Messing's lab
 The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
 University
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
 Tel: 732 445 3801
 Fax: 732 445 5735
 Email: bharti@waksman.rutgers.edu
 Seq primer: SP6
 Class: BAC ends
 High quality sequence start: 450.
 Location/Qualifiers

FEATURES
 source
 1..778
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /culivar="B73"
 /db_xref="taxon:4577"
 /clone="ZM5BHC0547L03"
 /lab_host="E. coli DH10B"
 /clone_id="ZM5BHC"
 /note="Vector: pTARBAC1.3; Site_1: BamHI; Site_2: BamHI"

ORIGIN
 Query Match 19.7%; Score 36.8; DB 9; Length 778;
 Best Local Similarity 34.8%; Pred. No. 23;
 Matches 39; Conservative 26; Mismatches 47; Indels 0; Gaps 0;

Oy 15 UGAUUAAGUAGAAAUUCCUUAUUAUUUUUUAUACUGCUACAUUUUUAAGAC 74
 Db 676 TGAATTCACAAATTAATATCAATTAACAAGTTTGAAGAACTCTTGAGATATACAC 617
 Oy 75 CCUAGUUAUUUACUUACCGCCAGAGGGGUGCAGCGUUCUCCAUA 126
 Db 616 TTTATTACAGATGTTTATCATCAAGATCGTTGAAAAAATCAAAAAATA 565

RESULT 7
 CL043477
 LOCUS
 DEFINITION CH216-57L9.RM.1 CH216 Xenopus tropicalis genomic clone CH216-57L9,
 genomic survey sequence.
 ACCESSION CL043477
 VERSION CL043477.1 GI:40499390
 KEYWORDS GSS.
 SOURCE Xenopus tropicalis (western clawed frog)
 ORGANISM Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 878)

AUTHORS
 Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
 Wards,E. and Wilson,R.
TITLE
 A physical map of the xenopus tropicalis genome
JOURNAL
 Unpublished (2003)
COMMENT
 Contact: Richard K Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@watson.wustl.edu
 Insert Length: 17500 Std Error: 0.00
 Seq primer: RM1 TACGACTCACTATGAGGAGA
 Class: BAC ends
 High quality sequence start: 4
 High quality sequence stop: 520.
 Location/Qualifiers

FEATURES
 source
 1..878
 /organism="Xenopus tropicalis"
 /mol_type="genomic DNA"
 /strain="Nigerian frog"
 /db_xref="taxon:8364"
 /clone="CH216-57L9"
 /sex="male"
 /cell_line="Stock 248 F7A2, inbred N7"
 /clone_id="CH216"
 /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
 BAC library"

ORIGIN
 Query Match 19.7%; Score 36.8; DB 10; Length 878;
 Best Local Similarity 29.8%; Pred. No. 23;
 Matches 31; Conservative 31; Mismatches 42; Indels 0; Gaps 0;

Oy 21 GAAGUAGAAGAAUUCUAGUUAUUAUUUUUUAUACUGCUACAUUUUUAAGCCCUUG 80
 Db 35 GAATTCACCAATGTATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 94
 Oy 81 UUAUUUAGCUUUAACGCCAGAGUGGUGCAGCCUUCUGCAA 124
 Db 95 TTAATGTTACTATACAGATGCAGATGCTGTTGCAAGATTTTAAAA 138

RESULT 8
 CN522916
 LOCUS
 DEFINITION clone GQ0123_E15 5', mRNA sequence.
 ACCESSION CN522916
 VERSION CN522916.1 GI:46841285
 KEYWORDS EST.
 SOURCE Populus trichocarpa x Populus deltoides
 ORGANISM Populus trichocarpa x Populus deltoides

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
 1 (bases 1 to 420)
 Morency,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A.,
 Retzetl,E. and Mackay,J.
 Arborea EST sequencing in Populus sp. (poplar)
 Unpublished (2004)
COMMENT
 Contact: John Mackay
 Centre de Recherche en Biologie Forestiere
 Universite Laval
 Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4
 Fax: 418 656 7493
 Email: jmackay@svs.ulaval.ca
 Center for Computational Genomics and Bioinformatics (CCGB)
 University of Minnesota, MN Id Identifier: MN5188653 Clone ID:
 GQ0123_E15 Clones available through: John Mackay, Ph. D. Professeur
 adjoint -Assistant professeur EMAIL: jmackay@svs.ulaval.ca Centre
 de Recherche en Biologie Forestiere (Forest Biology Research
 Center) Universite Laval Quebec, Quebec CANADA G1K 7P4
 Plate: 3 row: 15 column: E
 Seq primer: T3 Primer.
 Location/Qualifiers

RESULT	11
LOCUS	CO153454
DEFINITION	CO153454.1 486 bp mRNA linear EST 05-OCT-2004
ACCESSION	EN02010.5
VERSION	Exelixis FlyTag MN08 Bluescript Drosophila melanogaster cDNA clone EN02010 5, mRNA sequence.
KEYWORDS	CO153454
SOURCE	CO153454.1 GI:48907455
ORGANISM	Drosophila melanogaster (fruit fly)
REFERENCE	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 486) Nakanishi, M., Muzong, C., Peterson, E., Laufer, A., Leung, W., Platt, D. and Swimmer, C. Exelixis FlyTag EST Project MN08 Library Unpublished (2004) Contact: Stapleton, M.
TITLE	JOURNAL
COMMENT	BDGP Lawrence Berkeley National Lab One Cyclotron Rd, Berkeley, CA 94720, USA Fax: 510 486 6798 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu Plate: EN_20 row: A column: 10 High quality sequence stop: 412. Location/Qualifiers 1..486 /organism="Drosophila melanogaster" /mol_type="mRNA" /db_xref="taxon:7227" /clone="EN02010" /cell_line="mbn2" /clone_lib="Exelixis FlyTag MN08 Bluescript" /note="Vector: pBluescript, Site_1: NotI; Site_2: XhoI; oligodT primed from LPS induced mbn2 cell line."
ORIGIN	
Query Match	19.3%; Score 36; DB 7; Length 486;
Best Local Similarity	37.0%; Pred. No. 36;
Matches	40; Conservative 23; Mismatches 45; Indels 0; Gaps 0;
OY	53 AAUACUGCACAUUUUUAAGACCUCUAGUUUUAGCUUACCCGCCGAGUAGGGGUCA 112
Db	118 AGTAGTGCTATTAATTAAATGAAGTTACTATTAGTTTACTGTGCAGATGCTATTG 177
OY	113 GCUUUCCUGCAUUUVCAGGGCACUUGAGUUCACGCCUUGAUUUUAG 160
Db	178 GCAGCCCCATATAATCCAGACACCCCTCTCGCTTCCTTAATGATTAG 225
RESULT	12
LOCUS	CO153765
DEFINITION	CO153765 490 bp mRNA linear EST 05-OCT-2004
ACCESSION	EN02756.5
VERSION	Exelixis FlyTag MN08 Bluescript Drosophila melanogaster cDNA clone EN02756 5, mRNA sequence.
KEYWORDS	CO153765
SOURCE	CO153765.1 GI:4890766
ORGANISM	Drosophila melanogaster (fruit fly)
REFERENCE	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 490) Nakanishi, M., Muzong, C., Peterson, E., Laufer, A., Leung, W., Platt, D. and Swimmer, C. Exelixis FlyTag EST Project MN08 Library Unpublished (2004) Contact: Stapleton, M.
TITLE	JOURNAL
COMMENT	BDGP

	Lawrence Berkeley National Lab One Cyclotron Rd. Berkeley, CA 94720, USA
Fax:	510 486 6798
Email:	http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
Plate:	EN:27 row: E column: 8
High quality sequence stop: 401.	
FEATURES	
source	location/Qualifiers
	1..490
	/organism="Drosophila melanogaster"
	/mol_type="mRNA"
	/db_xref="taxon:7227"
	/clone="EN02756"
	/cell_line="mbn2"
	/clone_lib="Exelixis FlyTag MN08 Bluescript"
	/note=Vector: pBluescript; site_1: NotI; Site_2: XhoI;
	oligoDr primed from lps induced mbn2 cell line."
ORIGIN	
Query Match	19.3%; Score 36; DB 7; Length 490;
Best Local Similarity	37.0%; Pred. No. 36;
Matches	40; Conservative 23; Mismatches 45; Indels 0; Gaps 0;
Oy	53 AAUACUGCUCACAUUUUAAGACCUCUAGUAAUUUAGCUUUAACCGCCGAGAVGGSGUCA 112
Db	154 AGTAGTGATTATTAATAAATTGAAGTTAACCTTATTAGAATTTCGTGCAGATGCCATTATG 213
Oy	113 GCNUUCCUGCAUVUVCAGGAGGACCCUAGUGCACGCUUUGUUUUAG 160
Db	214 GCACCCCATTAAATCCAGACACCCCTCTGCTTCTTATATGATTAG 261
RESULT 13	
LOCUS	C0337794
DEFINITION	C0337794 539 bp mRNA linear EST 05-OCT-2004
VERSION	melanogaster cDNA clone EN15317 5, mRNA sequence.
KEYWORDS	C0337794.1 GI:49398069
SOURCE	EST.
ORGANISM	Drosophila melanogaster (fruit fly)
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 539) Nakanishi,M., Mazong,C., Peterson,E., Laufer,A., Leung,W., Platt,D. and Swimmer.C.
AUTHORS	Exelixis FlyTag EST Project MN08 Library
TITLE	Unpublished (2004)
JOURNAL	Contact: Stapleton, M.
COMMENT	BDCP
	Lawrence Berkeley National Lab One Cyclotron Rd, Berkeley, CA 94720, USA
	Fax: 510 486 6798
	Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
	Plate: EN:153 row: B column: 5
	High quality sequence stop: 474.
FEATURES	
source	Location/Qualifiers
	1..539
	/organism="Drosophila melanogaster"
	/mol_type="mRNA"
	/db_xref="taxon:7227"
	/clone="EN15317"
	/cell_line="mbn2"
	/clone_lib="Exelixis FlyTag MN08 Bluescript"
	/note=Vector: pBluescript; site_1: NotI; Site_2: XhoI;
	oligoDr primed from lps induced mbn2 cell line."
ORIGIN	
Query Match	19.3%; Score 36; DB 7; Length 539;
Best Local Similarity	37.0%; Pred. No. 36;
Matches	40; Conservative 23; Mismatches 45; Indels 0; Gaps 0;

Oy	53	AUACGCGUACAUUUUUAAGAACCCCUAUAUUUUUGAUUUUACGCCGCAAGUUGGGUGUCA	112
Dd	137	AGTAGTGTCTATTTTAATAATTAAAGTTAACTACTATTGATTGTTTACTGTTCAGATGCCTATTG	196
Oy	113	GCGUUCUGCAUAUACCAGGCGCAGCCUAGUGCACCCUUGAUUUUUAG	160
Dd	197	GCAGCCCCATATATATTCAGACACCCCTCTGCTTCTTATATGATTAG	244
<hr/>			
RESULT 14	B2344795	638 bp DNA linear GSS 12-NOV-2002	
LOCUS	B2344795	hg90h10.b1 WGS-SpicoLoRF (Jm107 adapted methyl filtered) Sorghum	
DEFINITION	bicolor genomic clone hg90h10 5', genomic survey sequence.		
ACCESSION	B2344795		
VERSION	B2344795.1 GI:24902202		
KEYWORDS	GSS.		
SOURCE	Sorghum bicolor (Sorghum)		
ORGANISM	Sorghum bicolor		
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoidae; Andropogoneae; Sorghum.		
AUTHORS	Radinowicz,P.D., O'Shaughnessy,A.L., Ballia,V., Dedhia,N., Ketzemburger,F., King,L., Miller,B., Muller,S., Mascimetto,L., Zutavern,I., Palmer,L., McCombie,W.R. and Martensen,R.A. Genomic shotgun sequences from Sorghum bicolor (methyl-filtered) Unpublished (2002)		
TITLE	Contact: W. Richard McCombie		
JOURNAL	Lita Annenberg Hazen Genome Sequencing Center Cold Spring Harbor Laboratory PO Box 100, Cold Spring Harbor, NY 11724, USA		
COMMENT	Tel: 516 367 8884 Fax: 516 367 8874 Email: mcombie@cshl.org Plate: hg90 row: h column: 10 Seq primer: -21M13univFwd Class: shotgun High quality sequence stop: 638. Location/Qualifiers		
FEATURES	source		
source	1..638		
	/organism="Sorghum bicolor"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:4558"		
	/clone="hg90h10"		
	/lab_host="Jm107 or DHSa"		
	/clone_lib="WGS-SpicoLoRF (Jm107 adapted methyl filtered)"		
	/note="Site 1: Xba I; Site 2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (x/y reads in M13mp19, .b/g reads in pUC19). The same ligation was transformed in either Jm107 or DHSa."		
ORIGIN			
Query Match	19.1%; Score 35.8; DB 9; Length 638;		
Best Local Similarity	34.5%; Pred. No. 42;		
Matches	41; Conservative 26; Mismatches 52; Indels 0; Gaps 0;		
Oy	11 GAUCGUAUUAGAGAUAUAUUUCGUUAUAUAUUUUUAUACUGCAUUUUUA	70	
Dd	215 GGATNGTTTAAACCTACTCTAATTTTTTAAATCATTTTTTAAACCTAGTGCTTCAA	274	
Oy	71 AGACCCUUAUUUUUAGCUUUACGCCGCAAGUGGGUGCAGCGUUCUCAUAUCC	129	
Dd	275 AGGCCCTTAAGTGGCCAATATATATACACCCGCTTGGGTCCAGCACTCAGCATCATGC	333	
<hr/>			
RESULT 15	BH876088	666 bp DNA linear GSS 05-AUG-2002	
LOCUS	BH876088		

DEFINITION	hr28h10_b1 WGS-Zmaysef (JM107 adapted methyl filtered) Zea mays
ACCESSION	BH876088
VERSION	BH876088.1
KEYWORDS	GSS.
SOURCE	Zea mays
ORGANISM	Zea mays
REFERENCE	Bukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 666)
AUTHORS	Rabinowicz,P.D., O'Shaughnessy,A.L., Ballifa,V., Dedhia,N., Ketzemburger,F., King,L., Miller,B., Muller,S., Nasimento,L., Zlatavern,T., McCombie,W.R. and Martienssen,R.A.
TITLE	Genomic shotgun sequences from Zea mays (methyl-filtered)
JOURNAL	Unpublished (2002)
COMMENT	Contact: W. Richard McCombie Iita Amnenberg Hazen Genome Sequencing Center Cold Spring Harbor Laboratory PO Box 100, Cold Spring Harbor, NY 11724, USA Tel.: 516 367 8884 Fax: 516 367 8874 Email: mccombie@cshl.org Plate: hr28 row: h column: 10 Seq primer: -2M13UniVfwd Class: Shotgun High quality sequence stop: 666. Location/Qualifiers
source	1..666
	/organism="Zea mays"
	/mol_type="genomic DNA"
	/culivar="B73"
	/db_xref="taxon:4577"
	/clone="hr28h10"
	/lab_host="JM107 or DH5a"
	/clone_lib="WGS-Zmaysef (JM107 adapted methyl filtered)"
	/note="Organ: Immature ears; Site_1: Xba I; Site_2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector l(x,y reads in M13mp19, b/g reads in pUC19). The same ligation was transformed in either JM107 or DH5a."
ORIGIN	
Query Match	19.1%; Score 35.8; DB 9; Length 666;
Best Local Similarity	34.5%; Pred. No. 42;
Matches	41; Conservative 26; Mismatches 52; Indels 0; Gaps 0;
CY	11 GAUCUGAUUAGAAGUAAGAUAUUCCUAGUUAUAUUVUUUAUACUGCACAUAUUUA 70 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : Db 223 GGATGTTTAAACTCTCAATTATTTTAATAATCATTTTTTAATAAAGTAGTGTCACCA 282 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : CY 71 AGACCUUAUUAUUUAUAGCUUACCGGCCGAGAUGGGUGCACGCGUUCUGCAUAUVC 129 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : Db 283 AGCCCTTAAGTGCACATATATATATACACC CGCTTGCGTCCAGCAACTCAGCATATGC 341

Search completed: December 22, 2005, 15:20:55

Job time : 1960.11 secs

This Page Blank (uspto)

/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 13291
/ LENGTH: 18008
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-13291

Query Match 17.5%; Score 32.8; DB 3; Length 18008;
Best Local Similarity 36.0%; Pred. No. 3.6;
Matches 36; Conservative 22; Mismatches 42; Indels 0; Gaps 0;

Qy 16 GAUUGAAGUAGAAUUCUAGUUAUUAUUUUUAUACUGCAUUUUUAAGAC 75
Db 6058 GACTCCATGTCATAAATAATTAATTAATTAATTAATTAATTAATTAATTTT 5999
Qy 76 CUAGUUAUUUAGCUUUAACGCCGAGUGGUGGACGCG 115
Db 5998 AGACGAGCTGCTGCTGCGCCAGCGTGGAGTGCACTG 5959

RESULT 7
US-09-949-016-16484
/ Sequence 16484, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: C1001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 16484
/ LENGTH: 26611
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-16484

Query Match 17.5%; Score 32.8; DB 3; Length 26611;
Best Local Similarity 32.4%; Pred. No. 4.2;
Matches 35; Conservative 26; Mismatches 47; Indels 0; Gaps 0;

Qy 12 AUCUGAUUAGUAGAAUUCUAGUUAUUAUUUUUAUACUGCAUUUUUA 71
Db 17672 ATGACATTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTTT 17731
Qy 72 GACCCUUAUUUAGCUUUAACGCCGAGUGGUGGACGCUUCC 119
Db 17732 TAGAGACAGAGTTTCGCTCTGTCAACCAGGCTGAGTGCACTGAGCC 17779

RESULT 8
US-09-949-016-13563
/ Sequence 13563, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: C1001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14

/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 13563
/ LENGTH: 50263
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-13563

Query Match 17.3%; Score 32.4; DB 3; Length 50263;
Best Local Similarity 36.6%; Pred. No. 7.2;
Matches 30; Conservative 21; Mismatches 31; Indels 0; Gaps 0;

Qy 1 GAAUUGUGUAGUCUGAUUAGAAUUCUAGUUAUUAUUUUUAUACUGC 60
Db 29101 GGAGATTTTGATATGGAATGGAAGAAATTTGTTGCAATGTTCTTAACACG 29160
Qy 61 UACAUUUUAAGACCUUAGU 82
Db 29161 CAGTGCTGTACGAAGCTTATTT 29182

RESULT 9
US-09-513-999C-9678
/ Sequence 9678, Application US/09513999C
/ Patent No. 6783961
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Duclert, A.Y.
/ TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
/ Patent No. 6783961
/ FILE REFERENCE: 59,US2,REG
/ CURRENT APPLICATION NUMBER: US/09/513,999C
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/122,487
/ PRIOR FILING DATE: 1999-02-26
/ NUMBER OF SEQ ID NOS: 36681
/ SOFTWARE: Patent.pm
/ SEQ ID NO 9678
/ LENGTH: 187
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 95
/ OTHER INFORMATION: m=a or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 122
/ OTHER INFORMATION: m=a or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 123
/ OTHER INFORMATION: y=c or t
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 127
/ OTHER INFORMATION: s=g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 130
/ OTHER INFORMATION: y=c or t
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 161
/ OTHER INFORMATION: m=a or c
/ FEATURE:


```

; NAME/KEY: misc_feature
; LOCATION: 183
; OTHER INFORMATION: w=a or t
US-09-513-999C-9678

Query Match          17.2%; Score 32.2; DB 3; Length 187;
Best Local Similarity 27.1%; Pred. No. 0.91;
Matches 35; Conservative 35; Mismatches 59; Indels 0; Gaps 0;

QY 37 UAGUUAUUUUUUUAUACUGCAUACUUUUUAGACCCUUAUUUUUAGCCUUAACCG 96
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 56 TGGTATATATATATATATATATATATATATATATATATATATATATATATAGCTTGCTTGTTG 115
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 97 CCCAGAGAGGGGUGCAGCGCUCUCCCAUAUUVACAGGACCAUAGUGUACGCCUUGUAGU 156
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 116 CCCAGAMTGGSTTYCAATATCCTTGCCCTCAGTGATCTCCCGGTCAGCTTCCCAAT 175
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 157 UUAGUGAC 165
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 176 TGCTGGGWC 184

RESULT 10
US-09-949-016-27423
; Sequence 27423, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 27423
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-27423

Query Match          17.2%; Score 32.2; DB 3; Length 601;
Best Local Similarity 29.6%; Pred. No. 1.4;
Matches 37; Conservative 30; Mismatches 58; Indels 0; Gaps 0;

QY 1 GAAAAGUGUGAUCGUAUUGAAGAAUUAUCCUAGUUAUUUUUUAUUAUUCGC 60
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 439 GTAAGTTTGCTGTTCACTCAATCTTCAATTTATATATATATATATTAAGATGATCATCTAT 498
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 61 UACAUUUUUUAGACCCUUAUUUUUAGCUUUAACGCCAGAGUGGGGUGCAGCGUCCU 120
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 499 TTCCTATATATCTTTATATATATACATAGATCTGGGGGCGGGTGGCTGCTCATGCT 558
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 121 GCAU 125
   |||:
Db 559 GTAAT 563

RESULT 11
US-09-949-016-27424
; Sequence 27424, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307

```

[illegible]

This Page Blank (uspto)


```

1 FILE REFERENCE: 108827.135
2 CURRENT APPLICATION NUMBER: US/09/925.065A
3 CURRENT FILING DATE: 2001-08-08
4 PRIOR APPLICATION NUMBER: US 60/243,096
5 PRIOR FILING DATE: 2000-10-24
6 PRIOR APPLICATION NUMBER: US 60/252,147
7 PRIOR FILING DATE: 2000-11-20
8 PRIOR APPLICATION NUMBER: US 60/250,092
9 PRIOR FILING DATE: 2000-11-30
10 PRIOR APPLICATION NUMBER: US 60/261,766
11 PRIOR FILING DATE: 2001-01-16
12 PRIOR APPLICATION NUMBER: US 60/289,846
13 PRIOR FILING DATE: 2001-05-09
14 NUMBER OF SEQ ID NOS: 957086
15 SOFTWARE: FastSeq for Windows Version 4.0
16 SEQ ID NO 267570
17 LENGTH: 600
18 TYPE: DNA
19 ORGANISM: Homo sapiens
20 US-09-925-065A-267570
21
22 Query Match 19.0%; Score 35.6; DB 4; Length 600;
23 Best Local Similarity 33.3%; Pred. No. 8.5;
24 Matches 26; Conservative 26; Mismatches 30; Indels 0; Gaps 0;
25
26 11 GAUCUGAUUAGAGAAUATTCUAGUUAUAUUUUUAUACUGCUACAUUUUUA 70

```

```

Oy      71 AGACCCUAGUUAAUUAUGCUUAC 94
||| ||| : ::||:: ::|
Db      226 AGACCCTCATTTATTTCTGTTTC 249

RESULT 8
US-09-925-065A-835698
; Sequence 835698, Application US/09925065A
; Publication No. US20050228172A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925_065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 835698
LENGTH: 538
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-835698

Query Match          18.6%; Score 34.8; DB 4; Length 538;
Matches Similarity 33.8%; Pred.No.14;
Match     27; Conservative 25; Mismatches 28; Indels 0; Gaps 0;

Oy      8 UGUAUCUCGAUUAAGAUAAGAAAUAUUCUAGUUAAUUAUUUUUUAUAUACUGCACAUTU 67
:::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      10 TGTGAACCTATTGTAATTTTAAAAATTTCAATTATATAATAATTTTATTTTCCTTCT 69
:::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```



```
US-09-925-065A-708859/C
? Sequence 708859, Application US/09925065A
? Publication No. US20050228172A9
? GENERAL INFORMATION:
?   APPLICANT: Wang, David G.
?   TITLE OF INVENTION: Identification and Mapping of Single
?   TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
?   FILE REFERENCE: 108827.135
?   CURRENT APPLICATION NUMBER: US/09/925,065A
?   CURRENT FILING DATE: 2001-08-08
?   PRIOR APPLICATION NUMBER: US 60/243,096
?   PRIOR FILING DATE: 2000-10-24
?   PRIOR APPLICATION NUMBER: US 60/252,147
?   PRIOR FILING DATE: 2000-11-20
?   PRIOR APPLICATION NUMBER: US 60/250,092
?   PRIOR FILING DATE: 2000-11-30
?   PRIOR APPLICATION NUMBER: US 60/261,766
?   PRIOR FILING DATE: 2001-01-16
?   PRIOR APPLICATION NUMBER: US 60/289,846
?   PRIOR FILING DATE: 2001-05-09
?   NUMBER OF SEQ ID NOS: 957086
?   SOFTWARE: FastSeq for Windows Version 4.0
?   SEQ ID NO 708859
?   LENGTH: 1265
?   TYPE: DNA
?   ORGANISM: Homo sapiens
US-09-925-065A-708859
```

	Query Match	Similarity	Score	DB	Length
Beet Local	32.0%	Pred.	No. 28;		
Matches	33;	Conservative	27;	Mismatches	43; Indels 0; Gaps 0
Qy	10	UGAUCGUAUAGAAACUAAAUUCCUAGUUAAUUUUUUUAUACUCGACAUUUUU	69		
Db	730	TTATTGGATCTTAATAAATAATTAATAATTGATTTTTTATTTTTAAAAATATTACTTTTT	671		
Qy	70	AAGACCUCUAGUUUUUAGCUUACCGCCAGAGUGGUGCA	112		
Db	670	TAGAACAAGGCTTGTCCTGTGTGTTGCACAGCTGAATGCA	628		

```

RESULT 14
US-09-925-065A-88724
; Sequence 88724, Application US/09925065A
; Publication No. US20050228172A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.115
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 88724
LENGTH: 733
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-88724

```

Query Match	18.2%;	Score 34;	DB 4;	Length 733;
Best Local Similarity	36.1%;	Pred. No. 26;		
Matches	39;	Conservative	23;	Mismatches 46;
				Indels 0;
				Gaps 0

Qy 12 AUCGUAUUAAGUAGAAAUUCCUAGUUUAUUUUUUUAUUCUGACAUUUUUUA 71
 19 ATACGAGTAGAATTTTAAATAATGACATCTATATATAAATGTACTTTTTCCTTTT 78
 Qy 72 GACCCUUAUUUAUAGCUUACCCGCCAGAGAGGGGUGACGGUUC 119
 Db 79 TTTCGACGAGAGTCTCGCTCTGTCCGCCAGCTGAGAGTGAAGTGGAGTC 126

```

RESULT 15
US-10-262-552-33
; Sequence 33, Application US/10262552
; Publication No. US20030125289A1
; GENERAL INFORMATION:
; APPLICANT: Gelb, Bruce D.
; APPLICANT: Tartaglia, Marco
; TITLE OF INVENTION: NOONAN SYNDROME GENE
; FILE REFERENCE: 2420/13859-US1
; CURRENT APPLICATION NUMBER: US/10/262,552
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 60/326,532
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 300000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(300000)
; OTHER INFORMATION: where n may be a or g or c or t/u, unknown or other
US-10-262-552-33

```

Query Match	18.1%	Score 33.8	DB 6	Length 300000
Best Local Similarity	37.7%	Pred. No. 2.6e+02		
Matches	26	Conservative	21	Mismatches 22; Indels 0; Gaps 0;
QY	51	UUAAUAACGUCNA	CATUUUUAAGCCCUAAGUUUAGCGUUGCGCAGAUAGGGUG	110
	:::::	:: :	: :::	: :::
Db	293826	TTAATTCTGTAAATTTTAAAAAGTTCTTTTGTTCACTGTGTGCCCAAGCTGAAGTG		293885
QY	111	CAGCGUUC	119	
	:	:: :	:	
Db	293886	CAGTGTAC	293894	

Search completed: December 22, 2005, 16:20:59
Job time : 438.281 secs

This Page Blank (uspto)

Result No.	Score	Query Length	DB	ID	Description
C 1	33.6	18.0	176503	7	US-11-121-086-53
C 2	32.6	17.4	1310	6	US-10-750-185-54747
C 3	32.6	17.4	101046	6	US-10-995-561-13330
C 4	32.4	17.3	1185	6	US-10-750-185-134828
C 5	31.8	17.0	164810	7	US-11-121-086-6
C 6	31.8	17.0	172543	7	US-11-121-086-6
C 7	31.4	16.8	110000	7	US-11-155-492-1
C 8	31.4	16.8	119036	6	US-10-995-561-13314
C 9	31	16.6	879	6	US-10-750-185-36820
C 10	31	16.6	990	6	US-10-750-185-42447
C 11	30.8	16.5	4244	6	US-10-753-626-4256
C 12	30.8	16.5	5935	7	US-11-177-987-42
C 13	30.6	16.4	201	6	US-10-995-561-53218
C 14	30.6	16.4	110950	6	US-10-857-780-2
C 15	30.6	16.4	403278	6	US-10-995-561-13421
C 16	30.4	16.3	936	6	US-10-750-185-53040
C 17	30	16.0	1926	6	US-10-750-185-53040
C 18	30	16.0	2088	6	US-10-750-185-57685
C 19	29.8	15.9	201	6	US-10-995-561-38112
C 20	29.8	15.9	201	6	US-10-995-561-38112
C 21	29.8	15.9	3042	6	US-10-750-185-53776
C 22	29.8	15.9	319608	7	US-11-145-703-1
C 23	29.6	15.8	1637	6	US-10-750-185-11849

C 24	29.6	15.8	143203	7	US-11-121-066-42	Sequence 42, Appl
C 25	29.6	15.8	398287	6	US-10-995-561-13396	Sequence 13396, A
C 26	29.6	15.8	1089214	7	US-11-117-187-211	Sequence 211, Ap
C 27	29.4	15.7	2261	6	US-10-750-185-58700	Sequence 58700, A
C 28	29.4	15.7	137935	6	US-10-995-561-13278	Sequence 13278, A
C 29	29.4	15.7	197781	7	US-11-112-908-34	Sequence 34, Appl
C 30	29.4	15.6	1128000	6	US-10-995-561-13266	Sequence 13266, A
C 31	29.2	15.6	450	6	US-10-750-185-20090	Sequence 20090, A
C 32	29.2	15.6	843	6	US-10-750-185-38003	Sequence 38003, A
C 33	29.2	15.6	7445	7	US-11-177-987-8	Sequence 8, Appl
C 34	29.2	15.6	187745	7	US-11-121-086-83	Sequence 83, Appl
C 35	29	15.5	881	6	US-10-750-185-57774	Sequence 57774, A
C 36	29	15.5	110000	7	US-11-155-492-1	Sequence 1, Appl
C 37	29	15.5	197096	7	US-11-121-068-107	Sequence 107, Appl
C 38	28.8	15.4	201	6	US-10-995-561-70950	Sequence 70950, A
C 39	28.8	15.4	435	7	US-11-194-246-475	Sequence 475, App
C 40	28.8	15.4	600	6	US-10-750-185-21810	Sequence 21810, A
C 41	28.8	15.4	706	6	US-10-986-501-11	Sequence 11, Appl
C 42	28.8	15.4	979	6	US-10-750-185-47956	Sequence 47956, A
C 43	28.8	15.4	1280	6	US-10-750-185-53636	Sequence 53636, A
C 44	28.8	15.4	1459	6	US-10-750-185-30881	Sequence 30881, A
C 45	28.8	15.4	1692	6	US-10-750-185-47673	Sequence 47673, A

ALIGNMENTS

```

RESULT 1
US-11-121-086-53/c
Sequence 53, Application US/11121086
Publication No. US2005026459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S
APPLICANT: NIELSEN, KIRSTEN V.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09318, 6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
PRIORITY FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIORITY FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 53
LENGTH: 176503
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-53

```

Query Match	Score	DB 7	Length
18.0%	33.6		176503

Best Local Similarity 26.8%; Pred. No. 4.4;

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040

Db 28101 ATTTAAATTCATCTTCTAATTGTCTTTTGGGTTTGTGTTGTTCTTA 28042

77 UGAGUUAUAGCUUACCGCCAGGAUGGGUGCAGCGUCCUGCAUAUC 128

Db 28041 TTTGTTTTAGCTCTGTTGCCAGGCTGGAGTGCAGTGCCACAATCATAGC 27990

RESULT 2

; Sequence 54747, Application US/10750185

GENERAL INFORMATION:

APPLICANT: DENISE, Sue K.

APPLICANT: ROSENFELD, David

APPLICANT: BATES, Stephen


```

: TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
: FILE REFERENCE: MM1100-2
: CURRENT APPLICATION NUMBER: US/10/750,185
: CURRENT FILING DATE: 2003-12-31
: PRIOR APPLICATION NUMBER: US 60/437,482
: PRIOR FILING DATE: 2002-12-31
: NUMBER OF SEQ ID NOS: 64922
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 54747
: LENGTH: 1310
: TYPE: DNA
: ORGANISM: Bovine 19866880573139
: US-10-750-185-54747

```

Query Match	17.4%	Score 32.6;	DB 6;	Length 1310;
Best Local Similarity	27.8%	Pred. No. 1.1;		
Matches 42;	Conservative 35;	Mismatches 74;	Indels 0;	Gaps 0;

[illegible]

Oy 89 CUUUAACCGCCAGAGUAGGSGUGCAGCGUUCUGCAUAUACAGGAGCACCUAAGUGCAGCC 148
 ||| : |||
 Db 552 AATCAGTTTTCAGTTTTCACAACACTTGCTTGAAATCATGCTCTTCTAATCAAGGTTATAAT 493
 ||| : |||

Qy	Db
149	uugUAGuuuvAGUGGACuuVAGGCGAAAGAA 179
492	TTGTAGCTTAATACTGCTTTTGGCTAGAAAA 462

RESULT 3
US-10-995-561-13330/C

Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OR INVENTION: CHEMICAL COMPOSITIONS FOR

METHODS OF DRUG RESPONSE,
CARDIOVASCULAR DISORDERS AND
TITLE OF INVENTION:
TITLE OF INVENTION:
DETECTION AND USES THEREOF
FILE REFERENCE: CLO01559
SUBMITTER ABBREVIATION NUMBER: US/70/866 E-1

```

; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO. 13330

```

```

?
? LENGTH: 101046
? TYPE: DNA
? ORGANISM: Homo sapiens
FEATURE:

```

```

NAME/KEY: mbc_feature
LOCATION: (1)...(101046)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see
SI-10-995-561-1330

```

Query Match	17.4%	Score 32.6;	DB 6;	Length 101046;
Best Local Similarity	32.4%;	Pred. No. 7.1;		
Matches 36:	Conservative	26;	Mismatches	49;
			Indels	0;
			Gaps	0;

[illegible]

69 UAA GACCCUUA GUUUAUUGACUUAUACCCGCCAGAGUAGGUGCAGCCUUC 119
 20925 TTA TTTT GAGACGAGTCTC ACTCTGTCA CCCCAGCTGTGATGTCAGTGGGC 20875

RESULT 4
US-10-750-185-34828/C

Sequence 34828, Application US/10750185
Publication NO. US20050260603A1
GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.
 ; APPLICANT: DENISE, Sue K.
 ; APPLICANT: KERR, Richard
 ; APPLICANT: ROSENFELD, David
 ; APPLICANT: HOLM, Tom
 ; APPLICANT: BATES, Stephen
 ; APPLICANT: FANTIN, Dennis
 ; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

```

; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
;

```

```
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34828
; SEQ ID NO 34829
```

```

; TYPE: DNA
; ORGANISM: Bovine 19866880730580
US-10-750-185-34828

```

Query Match	17.3%;	Score 32.4;	DB 6;	length 11
Best Local Similarity	31.2%;	Pred. No. 1.2;		
Matches 53;	Conservative 31;	Mismatches 86;	Indels	

```

QY      17 AUAAGAAAGUAGAAGAAAUUCCUGAUGUAUAUAUUUUUUUAUACUGGCUACAUU
           ||| ||| :||| | | | | | | | | | | | | | | | | | | | | | |
DB      985 ATTGAATGCATTGAAATCCTTAAGAAAAAGCTACATGACCAGAAGTTTCAT

```

925 ACGGGTTTACCAATTCCTCTGAGATATTATGGCAGTATTCACAAATT

Db 865 GTACATTCGATGTCTTACTTTAATCTACCATGCACCGAGAATTAAAT

RESULT 5
US-11-121-086-4/c
: Sequence 4, Application US/11121086

```

; PUBLICATION NO.: 05200502004359A1
;
; GENERAL INFORMATION:
;
; APPLICANT: POULSEN, TIM S.
;
; APPLICANT: NIELSEN, KIRSTEN V.

```

TITLE OF INVENTION: NOBLEIC ACID PROBES AND NOBLEIC ACID ANAL
 FILE REFERENCE: 09138.6000-0000
 CURRENT APPLICATION NUMBER: US/11/121,086
 CURRENT FILING DATE: 2005-05-04

```

; RECORD INFORMATION NUMBER: 00/00/0000
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3

```

```

; LENGTH: 164810
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```
Query Match      17.0%; Score 31.8; DB 7; Length 16
Best Local Similarity 36.4%; Pred. No. 15;
```

17 AUUAGAUAAGAAAUUCUAGUUAUAUAUUUUUAUACUGCUACAUT

77 UUAGUUAUUUAGCUTUUAACCCAGGAGGAGGUGCAGCG 115

RESULT 6


```

US-11-121-086-6/c
; Sequence 6, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: US/11/121, 086
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 172543
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-6

Query Match      17.0%; Score 31.8; DB 7; Length 172543;
Best Local Similarity 36.4%; Pred. No. 16;
Matches 36; Conservative 21; Mismatches 42; Indels 0; Gaps 0;

Qy      17  AUAAGAUAAGAAAUUUCUAGUUUAUAUUAUUUAUUAAGACCC 76
Db      169286  ATAAATAATCTCAATATCAATTTTAAATAATTAATCTTCTCAATTTTITGA 169227
Qy      77  UUAGUUAUUUAGCUUUAUACCGCCAGAUUGGGUGCAGC 115
Db      169226  AAAATTGTCTCGCTGTGTTGCCCAAGCTGGAGTGTAGTG 169188

RESULT 7
US-11-155-492-1
; Sequence 1, Application US/11155492
; Publication No. US20050266479A1
; GENERAL INFORMATION:
; APPLICANT: Weissenbach, Jean
; APPLICANT: Hazen, Jam11
; TITLE OF INVENTION: CLONING, EXPRESSION AND CHARACTERIZATION OF THE SPG4
; TITLE OF INVENTION: GENE RESPONSIBLE FOR THE MOST COMMON FORM OF AUTOSOMAL
; TITLE OF INVENTION: DOMINANT SPASTIC PARAPLEGIA
; FILE REFERENCE: R-341894
; CURRENT APPLICATION NUMBER: US/11/155,492
; CURRENT FILING DATE: 2005-06-20
; PRIOR APPLICATION NUMBER: US/09/830,902
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: FR 99 11097
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: PCT/FR00/02433
; PRIOR FILING DATE: 2000-09-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 1
; LENGTH: 110000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (9932)...(10471)
; FEATURE:
; NAME/KEY: Intron
; LOCATION: (10472)...(33718)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (33719)...(33805)
; FEATURE:
; NAME/KEY: Intron
; LOCATION: (33806)...(35748)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (35749)...(35832)

```

1	FEATURE:	intron	
2	NAME/KEY:	(35833) ... (45022)	
3	LOCATION:		
4	FEATURE:	exon	
5	NAME/KEY:	(45023) ... (45118)	
6	LOCATION:		
7	FEATURE:	intron	
8	NAME/KEY:	(45119) ... (60863)	
9	LOCATION:		
10	FEATURE:	exon	
11	NAME/KEY:	(60864) ... (61051)	
12	LOCATION:		
13	FEATURE:	intron	
14	NAME/KEY:	(61052) ... (61927)	
15	LOCATION:		
16	FEATURE:	exon	
17	NAME/KEY:	(61928) ... (62061)	
18	LOCATION:		
19	FEATURE:	intron	
20	NAME/KEY:	(62062) ... (62344)	
21	LOCATION:		
22	FEATURE:	exon	
23	NAME/KEY:	(62345) ... (62438)	
24	LOCATION:		
25	FEATURE:	intron	
26	NAME/KEY:	(62439) ... (73173)	
27	LOCATION:		
28	FEATURE:	exon	
29	NAME/KEY:	(73174) ... (73248)	
30	LOCATION:		
31	FEATURE:	intron	
32	NAME/KEY:	(73249) ... (74633)	
33	LOCATION:		
34	FEATURE:	exon	
35	NAME/KEY:	(74634) ... (74705)	
36	LOCATION:		
37	FEATURE:	intron	
38	NAME/KEY:	(74706) ... (82788)	
39	LOCATION:		
40	FEATURE:	exon	
41	NAME/KEY:	(82789) ... (82864)	
42	LOCATION:		
43	FEATURE:	intron	
44	NAME/KEY:	(82865) ... (83102)	
45	LOCATION:		
46	FEATURE:	exon	
47	NAME/KEY:	(83103) ... (83194)	
48	LOCATION:		
49	FEATURE:	intron	
50	NAME/KEY:	(83195) ... (83334)	
51	LOCATION:		
52	FEATURE:	exon	
53	NAME/KEY:	(83335) ... (83414)	
54	LOCATION:		
55	FEATURE:	intron	
56	NAME/KEY:	(83415) ... (88129)	
57	LOCATION:		
58	FEATURE:	exon	
59	NAME/KEY:	(88130) ... (88172)	
60	LOCATION:		
61	FEATURE:	intron	
62	NAME/KEY:	(88173) ... (89561)	
63	LOCATION:		
64	FEATURE:	exon	
65	NAME/KEY:	(89562) ... (89641)	
66	LOCATION:		
67	FEATURE:	intron	
68	NAME/KEY:	(89642) ... (91162)	
69	LOCATION:		
70	FEATURE:	exon	
71	NAME/KEY:	(91163) ... (91233)	
72	LOCATION:		
73	FEATURE:		

	Query Match	16.5%; Score 30.8; DB 7;	Length 5935;
	Best Local Similarity	32.9%; Pred. No. 7.5;	
	Matches	27; Conservative	23; Mismatches 32; Indels 0; Gaps 0;
Qy	10 UGAUCUGAUVAGAAUAAAACCUAGUUAAUAUUUUUUUAUAUCUGCACAATUUU	69 ::: ::: ::: ::: ::: TGTGATTATTTAAAGGGAGAATAAATTAACAAGCACAAGATTATCATTCATGCTTTAT	5803
Dd	5862 TGTGATTATTTAAAGGGAGAATAAATTAACAAGCACAAGATTATCATTCATGCTTTAT	5803	
QY	70 AAGACCCTUUNAGUUAUUUAGCU	91 : : ::: ::: :::	
Db	5802 TAAAGCCCTAAGATATCTTGTTTT	5781	

```
Query Match      16.4%; Score 30.6; DB 6; Length 110950;
Best Local Similarity 36.4%; Pred. No. 31;
Matches 28; Conservative 20; Mismatches 29; Indels 0; Gaps 0;
```

OY 11 GAUCUGAUUAGAAGUAGAAAAUUCGUAUUAAUUUUUUUAUACUGCUACAUUUUUA 70
 | | | | | :
DB 102820 GTTCGTGAAGAGAGAAAATAATTGATTTATGATAGCATTAATCAATCATTTTTC 102879

OY 71 AGACCCUUAGUUUUUA 87
 | | | | | : : : : : .

DB 102880 AGCCTCAAAGTAAATTTA 102896
 | | | | | : : : : :


```

RESULT 15
US-10-995-561-13421
; Sequence 13421, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13421
; LENGTH: 403278
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(403278)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13421

Query Match      16.4%; Score 30.6; DB 6; Length 403278;
Best Local Similarity 32.5%; Pred. No. 53;
Matches 38; Conservative 25; Mismatches 54; Indels 0; Gaps 0;

QY      16  GAVUAGAAGUAGAAAUUCCUAGUUAUUAUUUUUAUACUGUACAUUUUUAGACC 75
Db      194921 GATTAAGAAAAACAAAATATATGAGTTAAATTTTCTTTGTTTATGATC 194980

QY      76  CUUAGUUUUUAGCUUUUACCGCCGAGGAGUGGUGACGCGUCCUGCAUUAUCCAG 132
Db      194981 CTCAGATTCTTGAGATTAAATTCACAGCAGGATGACGCTCTTTGAGTTACATG 195037

```

Search completed: December 22, 2005, 16:41:00
Job time : 174.639 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 04:46:32 ; Search time 1490.97 seconds
(without alignments)
7205.664 Million cell updates/sec

Title: US-10-088-750C-3

Perfect score: 189

Sequence: 1 guagagugugaucuucuu.....uaggaauagaanaaacu 189

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl:
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_pro:*
10: gb_str:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	189	100.0	189	6	BD173513 Novel ter
2	189	100.0	189	6	BD177017 Novel tra
3	189	100.0	9264	13	AF014388 Drosophila
4	187.4	99.2	1345	6	AR496112 Sequence
5	187.4	99.2	1345	6	AR511394 Sequence
6	123.6	65.4	188	6	BD173514 Novel ter
7	123.6	65.4	188	6	BD177018 Novel tra
8	123.6	65.4	9185	13	AF218039 Cricket p
9	45.4	24.0	190	6	BD173516 Novel ter
10	45.4	24.0	190	6	BD177020 Novel tra
11	45.4	24.0	8550	13	AF183905 Black que
12	40.4	21.4	1347	13	CVU55052 Compox viru
13	40.4	21.4	224499	13	AF482758 Compox vi
14	40.4	21.4	334371	14	AC109571 Rattus no
15	40	21.2	175092	9	AC122491 Mus muscu
16	40	21.2	252108	9	AC138642 Mus muscu
17	39.8	21.1	53711	8	AB038985 Homo sapi
18	39.8	21.1	179216	14	AC074016 Homo sapi

19	39.8	21.1	181150	8	AC007256 Homo sapi
20	39.6	21.0	5059	6	C0597290 Sequence
21	39.6	21.0	52149	14	AC012958 Drosophila
22	39.6	21.0	185200	2	AC092401 Drosophila
23	39.6	21.0	191558	2	AC013431 Drosophila
24	39.6	21.0	303344	2	AE003500 Drosophila
25	38.4	20.3	108176	9	AL845327 Mouse DNA
26	38.2	20.2	9010	13	AF178440 Triatoma
27	38.2	20.2	139631	8	AC146534 Otolium
28	38	20.1	234660	14	AC134709 Rattus no
29	38	20.1	240475	14	AC099247 Rattus no
30	37.6	19.9	9812	13	AF536531 Aphid let
31	37.6	19.9	9829	6	AX763949 Sequence
32	37.6	19.9	110000	1	AE017332.6 Continuation (7 of
33	37.4	19.8	785	10	BV599101 S217P6180
34	37.4	19.8	174520	8	AC093432 Homo sapi
35	37.4	19.8	183127	14	AC150913 Bos tauru
36	37.4	19.8	227069	4	AC149669 Bos tauru
37	37.4	19.8	241957	14	AC119141 Rattus no
38	37.2	19.7	186045	5	AL935182 Zebrafish
39	37.2	19.7	231337	14	AC137758 Rattus no
40	37	19.6	218951	14	AC125915 Rattus no
41	36.8	19.5	110000	14	AC107134_2 Continuation (3 of
42	36.8	19.5	128674	14	AC098368 Rattus no
43	36.8	19.5	147586	14	CR735143 Danto rer
44	36.8	19.5	180401	8	AC007489 Homo sapi
45	36.8	19.5	212537	14	AC109658 Rattus no

ALIGNMENTS

RESULT 1	BD173513	189 bp	RNA	linear	PAT 18-FEB-2003
LOCUS	BD173513				
DEFINITION	Novel tertiary structure having ability to accelerate translation activity.				
ACCESSION	BD173513				
VERSION	BD173513.1	GI:28414844			
KEYWORDS	WO 02061080-A/3.				
SOURCE	Drosophila C virus				
ORGANISM	Drosophila C virus				
REFERENCE	Viruses; ssRNA positive-strand viruses, no DNA stage; Dicistroviridae; Crispavirus.				
AUTHORS	1 (bases 1 to 189)				
TITLE	Nakashima,N. and Kanamori,Y.				
JOURNAL	Novel tertiary structure having ability to accelerate translation activity				
Patent:	WO 02061080-A 3 08-AUG-2002;				
JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF					
AGRICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE					
FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA, YASUSHI KANAMORI					
COMMENT	OS Drosophila C virus				
PN	WO 02061080-A/3				
PD	08-AUG-2002				
PF	31-JAN-2001 WO 2001JP000641				
PR	25-JAN-2001 JP 01P 016746				
PI	NOBUHIKO NAKASHIMA, YASUSHI KANAMORI				
PC	C12N15/11,C12N15/86,C12P21/02				
CC	Novel tertiary structure having ability to accelerate translation activity				
FH	Key	Location/Qualifiers			
FT	source	1.189			
FEATURES	Location/Qualifiers	/organism='Drosophila C virus'.			
source	1.189				
ORIGIN	Location/Qualifiers	/organism='Drosophila C virus' /mol_type='genomic RNA' /db_xref='taxon:64279'			
Query Match	100.0%;	Score 189;	DB 6;	Length 189;	
Best local Similarity	64.0%;	Pred. No. 2.3e-35;			

	Matches	121, Conservative	68, Mismatches	0, Indels	0, Gaps	0, Ns
Qy	1	GUUAAAGVUGUAGUCUGUCUCCUUAUACAUAUUUGAGAGGUAAUAGAAGAGUAGUG	60			
Db	1	GTTAAGAGTGTATCTTGCTCCTTTACAAATTTTCAGAGGTTAAATMAAAGAGAGTAACTGTG	60			
Qy	61	CUAACUUAUUAUUGAGUUAACUAUUNUAGUUUAUCUGUCAGAGAGCCUUAUUGCAGGCC	120			
Db	61	CTATCTTAATTAATTAAGTTAACTAATTTAGTTTACTGTTCAGAGATGCTATATGGCAGGCC	120			
Qy	121	CAUAAUAVCCAGAGACCCUCUCUGUCUUUAUUGAUUAGUUGUGUCAUUAGAUAUAGA	180			
Db	121	CATATATATCCAGAGACACCTCTGCTCTTATATAGATTAGTGTCATTAGAAATAGA	180			
Qy	181	AAAUAAACCU	189			
Db	181	AAATTAACCT	189			

LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AF014388	Drosophila C virus strain EB, complete genome.	AF014388					
AF014388.1	GI:2388672						
	Drosophila C virus						
	Drosophila C virus						
	viruses; ssRNA positive-strand viruses, no DNA stage;						
	Dicistroviridae; Crispavirus.						
	1 (bases 1 to 9264)						
	Johnson, K.N. and Christian, P.D.						

RESULT 2				
Bd177017				
LOCUS	Bd177017	189 bp	RNA	linear
DEFINITION	Novel translatational activity-Promoting higher-order structure.			
ACCESSION	Bd177017			PAT 16-APR-2003
VERSION	Bd177017.1	GI:30014277		
KEYWORDS	JP 2002306168-A/3.			
SOURCE	<i>Drosophila C virus</i>			
ORGANISM	<i>Drosophila C virus</i>			

JOURNAL	undescribed virus family
PUBLISHED	J. Gen. Virol. 79 (Pt 1), 191-203 (1998)
REFERENCE	9460942
AUTHORS	2 (bases 1 to 9264)
TITLE	Johnson, K.N. and Christian, P.D.
JOURNAL	Direct Submission
FEATURES	Submitted (16-JUL-1997) Entomology, CSIRO, Clunies Ross St., Canberra, ACT 2602, Australia location/Qualifiers

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 189)	Nakaehime, N. and Kanamori, Y.	Novel transnational activity-promoting higher-order structure Patent : JP 2002306168-A 3 22-OCT-2002;	
	DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND ENTOMOLOGICAL HIROSHIMA PREFECTURAL INDUSTRIAL TECHNOLOGY PROMOTION ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND FISHERIES		
OS	<i>Drosophila</i> C virus		
PN	JP 2002306168-A/3		

```

CDS
    /organism="Hepatitis C virus"
    /mol_type="genomic RNA"
    /strain="EB"
    /db_xref="taxon:64279"
    799..6078
    /note="includes putative protease, helicase and RdRp
    domains; putative; ORF-1"
    /codon_start=1
    /product="replicase polypeptide"
    /protein_id="NC_05807.1"
    /db_xref="GI:2388673"

```

FEATURES	Location/Qualifiers
source	1..189
	/organism="Drosophila C virus"
	/mol_type="genomic RNA"
	/db_xref="taxon:64279"

DQVQKATPELFEVIRSCNTPQOHLMAHLHDKNFSAEILLYYTNDVNVLEISIT
 KODKKAANPELFEVIRSCNTPQOHLMAHLHDKNFSAEILLYYTNDVNVLEISIT
 PDDAFENRMGDMAVSCPKKEVGIETEKENGSGVTYTKLDSKDKTADLSVYEPK
 IYRDEKSDAGMIDSGSPDYEDPAKVCSKWEKAKSGSNMVKFLFEYAIRAIVGSE
 NSEYGCIDPVDIDIAFLQKGETLEIEEFYVADSPPMFTQYHFKSTIKRASKWQK
 DMDICLSDCKTYLAKKYEIKKILIAEHPILITLIGMIGVLSMAMTYWPSKSLDPE

[illegible]

VQSGSDVKTNRKQAIKREVEAGAEIYVTHALTQGGSDDAAHNMIMIDVOKATTNRSYTRG
 DREYDQNGNCTFVRGWSFLMPHYHVQAVFARLRPPNTIISLQSDMSLMOIPLSHPS
 ACDYVQLNDNVCVRLPFKNQDFRDCVMVNIASHMPCMRHRLVRLPILTSOQGLKSPF
 SAGMAMTFHYNMGLYKRYVVMVNLARCDCKIEITHEDGCEYEESISIQDCEYVNAV
 TRTDGGSIIGLYKYLKERKIIGMRHAGNDADEHGAECPLTQCELEAFALVANKK
 NISQSPVEYIEPMMWVBLGDSVPECKFVAGLGSIRFVQAVQVNIISPTRYGLSVPT
 MRPALPRTIILNNKYNHNLDSGLKCKGVDTVAISDVEYSASQDVCVRVMMNYNKNL
 KTKYQILTYIEBEAIRKQDDEFCMCALNRTTSPQPYAOKMRNAPGQAWMNGFEETDE
 TSNVYALAKDVEOBLIDECASQISNVILVDLKERDIRAKVNRKSTRVFSQSPHF
 VVAFQYFLPFAAMLMHNRISNEVAGTVNSDWMRIARLTKTKSHYVAGPFGND
 GSLVQILMAIEMEIVFWLKOFLIDENEGEKRIICGLMSHLVAHYIYEDNYM
 WHQSQGNBPFTVILNCLNYSIMRLSMIRNMEKBPRLKSMKMFNEVVALIYQGDV
 VANIADKAYEMFNQINISSVYTEMKREHYDEAKGTIDYVSKRLIEDYFLPKLRKRSPE
 LQTHVAAPLKIETVYENLNNRSIDPEDEILMSNITFAFREVYHGEKEYDKLSAVIA
 LKVPQELPENPQILTYIYNOYLRHIDIEYLAIDLVPF"
 <6267. .8972
 /note="ORF-2: putative: translation start site unknown"

JOURNAL	Patent: WO 02061080-A 4 08-ANG-2002; JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA, YASUSHI KANAMORI									
COMMENT	OS Cricket paralysis virus PN WO 02061080-A/4 PD 08-AUG-2002 PF 31-JAN-2001 WO 2001JP000641 PR 25-JAN-2001 JP 01P 016746 PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI PC C12N15/11, C12N15/86, C12P21/02 CC Novel tertiary structure having ability to accelerate CC translation activity									
FEATURES	FH Location/Qualifiers FT source 1. 188 FT /organism='Cricket paralysis virus'. Location/Qualifiers 1. 188 /organism="Cricket paralysis virus" /mol_type="genomic RNA" /db_xref="taxon:12136"									
ORIGIN										
Query Match	65.4%; Score 123.6; DB 6; Length 188;									
Beet Local Similarity	51.1%; Pred.No.1.4e-19;									
Matches	95; Conservative 52; Mismatches 39; Indels 0; Gaps 0									
OY	4 AAGAGUGAGUCUGUCUCUUAACAUUUUUAGAGAGUUAUAGAAGAGAAUGAGUGCAU 63									
Db	3 AAAATGTATCTTGGCTTGTAATTAACATTTTGGAGAGTTAAATTAACAAGTAGTGCTA 62									
OY	64 UCUUAUAUAUGAGUUAACUUAUUUUUUUACUGUCAGAGUCCUUAUUGAGAGCCCAU 123									
Db	63 TTTTGTATTATAGTTAGCTATTATTTAGCTTTTACGTTCCAGAGAGCTTAGTGAGCCAC 122									
OY	124 AAUAGCCGAGACACCCUCUCUGUCUCUUAUAUGAUUAGUGUGUCAUUAGAUAAGAAA 183									
Db	123 AATATCCAGAGAGCCCTCTCTCGGTTTTCAGATTAGTAGTGCAGAAAACTTAAGAAAT 182									
OY	184 UAACCU 189									
Db	183 TTACCT 188									
RESULT 7										
BD177018										
LOCUS	BD177018 188 bp RNA linear PAT 16-APR-2003									
ACCESSION	Novel translational activity-promoting higher-order structure.									
VERSION	BD177018.1 GI:30014278									
KEYWORDS	JP 2002306168-A/4.									
SOURCE	Cricket paralysis virus									
ORGANISM	Cricket paralysis virus Viruses; ssRNA positive-strand viruses, no DNA stage; Dicistroviridae; Crispavirus. 1 (bases 1 to 188)									
REFERENCE	Nakashima, N. and Kanamori, Y. Novel translational activity-promoting higher-order structure Patent: JP 2002306168-A 4 22-OCT-2002; DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND ENTOMOLOGICAL HIROSHIMA PREFECTURAL INDUSTRIAL TECHNOLOGY PROMOTION ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND FISHERIES									
AUTHORS	OS Cricket paralysis virus									
TITLE	PN JP 2002306168-A/4									
JOURNAL	PD 22-OCT-2002									
COMMENT	PF 25-JAN-2001 JP 2001016746 PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI PC C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02// PC (C12N15/09, C12R1:92), C12N15/00, C12N5/00, C12N15/00, C12R1:92) CC Novel translational activity-promoting higher-order structure FH Key Location/Qualifiers FT source 1. 188									

FEATURES	location/Qualifiers	FT
source	1..188	
	/organism="Cricket paralytic virus"	
	/mol_type="genomic RNA"	
	/db_xref="taxon:12136"	
Query Match	65.4%; Score 123.6; DB 6; Length 188;	
Best Local Similarity	51.1%; Pred. No. 1.4e-19;	
Matches	95; Conservative 52; Mismatches 39; Indels 0; Gaps 0;	
QY	4 AAGAGUGAUCUCUGCCUUAUAACAUAUUUGAGAGUUAUAGAAGAGUAGUGCUA 63	
DB	3 AAATGTGATCTTCTGTGTAATACAAATTTGAGAGGTTAATTAATTCAAGTAGTGCTA 62	
QY	64 UCUAUAUAUAGUUAUAUAUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUAUUUA 123	
DB	63 TTTTGTATTTAGGTATGCTATTTATTTAGTTTACGTTCCAGGATGCTGACGCCAC 122	
QY	124 AAUUAUCCAGACACCCUUCUGUCUUAUAUUAUAGUUAUUAGUUGUUAUUUAUAGAUA 183	
DB	123 AATATCCAGGAGGCCCTCTCTGCGGCTTTTCAGATTGCTAGTAGTCAAAAACCTTAAGAAAT 182	
QY	184 UAACCU 189	
DB	183 TTACCT 188	
RESULT 8		
LOCUS	AF218039 9185 bp RNA linear VRL 02-JUL-2000	
DEFINITION	Cricket paralytic virus nonstructural polyprotein and structural	
ACCESSION	AF218039	
VERSION	AF218039	
KEYWORDS	AF218039.1 GI:8895506	
SOURCE		
ORGANISM	Cricket paralytic virus	
	Cricket paralytic virus	
	viruses; ssRNA positive-strand viruses, no DNA stage;	
	Dicistroviridae; Crispavirus.	
REFERENCE	1 (bases 1 to 9185)	
AUTHORS	Wilson,J.E., Powell,M.J., Hoover,S.E. and Sarnow,P.	
TITLE	Naturally occurring dicistronic cricket paralytic virus RNA is	
JOURNAL	regulated by two internal ribosome entry sites	
PUBMED	Mol. Cell. Biol. 20 (14), 4990-4999 (2000)	
AUTHORS	2 (bases 1 to 9185)	
TITLE	Wilson,J.E., Powell,M.J., Hoover,S.E. and Sarnow,P.	
JOURNAL	Direct Submission	
PUBMED	Submitted (20-DEC-1999) Microbiology & Immunology, Stanford	
AUTHORS	University, 299 Campus Drive, Stanford, CA 94305, USA	
TITLE	Location/Qualifiers	
JOURNAL	1..9185	
PUBMED	/organism="Cricket paralytic virus"	
AUTHORS	/mol_type="genomic RNA"	
TITLE	/db_xref="taxon:12136"	
JOURNAL	709..6024	
PUBMED	/codon_start=1	
AUTHORS	/product="nonstructural polyprotein"	
TITLE	/protein_id="AAF80998.1"	
JOURNAL	/db_xref="GI:8895507"	
PUBMED	/translation="MSFQDTNNATNNINSLEELAAOELIAAOFESGNDLGFCTFYVO	
AUTHORS	SRPQLDLBSECTCHDDFDCCGCDRIKREELKLIPLISDVGVNFEENKGLVWKFQV	
TITLE	NCPENHRGSGTENGLLIVSPERFMDHLDFQOMKLVSSNDCRAFLRRTGLMSGD	
JOURNAL	VESNGAPQSRPVACNDPRAIRLEKALQREDEITSLIKLROEIKNNRIYTGSGF	
PUBMED	DLKAKXGQGLNGNLRIICDFLENSLPTLAQIQITVLITTDKVKIKEDILKVAI	
AUTHORS	LLVLRILAMVWKVRAALIVITLFWHFGFGRKOLIDLKDKLOTTTGAQRETL	
TITLE	EEVVHWPEDCGKILPAVIAFPALIKKIPGKODMNYISRLDRIPRAIGSKKIYVC	
JOURNAL	SEVPILSDEYKAVVLGKELKTQGLYDEIHMAKRIHYLDLDERNKITLPLETNAK	
PUBMED	VEDELTKGKTSSEKIPDRIDRLFTTMLFPKSLSYEOLLSVKGSGKRMPIYVWL	
AUTHORS	TGESGIRGKTMVPLCIDILRBMGIIVKPDVAYHQAIVARVETEMDGVNGQKIVIVDD	

CDS

AFOLKDKPKENPELFEVIRTCNTEPQHLHMAALODKMYSOAEVLTYTNOFOVLE
SITPPDAFYNMKTAAHVVOIKOEKSIIVNKRABEENALVTLNDKDEALDSVBER
OKMPPDESATKMIIDGEPISYDERPARICTKAMEBEKTHQLOMLEAVASRVVAC
GSETEYIDVDETFYSLNLSQGFPAAGSLIEMEBEPASDSTQNLLEYKNIPEKT
KMSKMTILDEOISALSTKIRELNKAKYISEHPYLLALGFIWMI SAFMYSPFER
TLTDJITTSVEVSGSDNKTQISKRVAEVSQGVTKTKTVAEVSQSDSTMYN
KITRVEVSGSDSKTOKORTKVEGKELKEAEOTGSCDAALVLDVQKNTYCL
YVEBMVNGEMKRYBLATPLRIGVCMMPHYRTITVAKYARPAATNLYFESQPNDDVI
VVPISHPIAPAEVEVELTACTRIHYDETRDVCVUNHMRMCHPRHDIKHYKVS
DQGLRVFGQSLATPHOSANELCAVOMLOAIRPDELTITHEIDMFPSEEST
ORDCEYNAFPQTGNCOSI VGLYNRMERKLIQMIIPONVSECGVACPLIOEALIMG
LNRLEKIDPVANNITVOCCEPSPDIKOTMSETEBEGKCAIGKSNIKGAVKTTLLK
SCIYMSKSPITRPAHLTRTLRPNGEI VDPIMKGLKCGVATAVLADEIVESALDVK
QVVLTOYNSMDANKYRRLTVEBATOGTDDDMKGIAROTSPBNGYRFOKRLDPK
QDMMSGEOCTPMTORABELRBDYBELDNCAGKI IKGVYVDTLKBERRIRBYDVG
KTRVFSAGPQHVVAFRKFLPFAVYLMNRI DNEIAGINVTSTDERINAKRLKAG
KSVIADGPNFDSLVAQFQSGCKSPYPMFKTFNDVNTDGRNMTI CIGLMTIYV
HSHVSYGDNVMTSHOPSQNPFTVILINLYNSMI MRVMTLARKLAPEKQMKKFR
ENVMESIYSGDNCLNISDRVVEWFOITISEOMKEIKHEYDECKTGMVKEPSELSEI
HFLKRRFVESHQLORTVAPLOKDVYEMELMNTNTIDPNEILMNTNIMTAPEIYYHCK
SEYOKLSGIEDLMMKGLPQOPQILTRFKATLMATMLADEVYVP"
<6217..8904
/codon_start=1
/product="structural polypeptide"
/protein_id="AAF80999.1"
/db_xref="GI:8895508"
/translation="ATPDDKOBNSHIEBEDKRLMSOEKLYHPVSEGTTSBETLAPDI
VNLSINYDMTRREDRIHSIXDPLSGPILITNLMSSSDVPEKOLYTNAPFEVLISNA
MYODKLKGFVGLRALTVKRVQVNSQPFQOGLMLOIYIPYQMPNRVTLINETLOGRS
GCPITDLELSTVEVEMRI PYVSPHLYNLTTCGGSFOSIVVYVYSLMDVQSTGSI
EYTWALHEVDVQVPTGANIFETGNSPWLFIARINAGDLETTEEMKRLMHTKYLR
PARIYAAAKELKOLETNSPSTALGOISELTLSHPIVAGNI ESPMNAISAAADL
AKLGFSPITQGTIGECKLAGCGKMANFDMDSHRAALSTNIEIKESBLATSID
EMDSLKSLSTPNYMDRFTWKTSDYNTVLMONVSPFKVRYSAITIDREFCTHMGV
ANATYMGSIYVTFKFKYQYHSGRLISPIPYVNTTISTGPDVSRQKTIIVDLR
TSTEVSPVPIYIARPMYLCIRPESMISKONKDALMNCVSGIVREVEANQVAAQ
NVFSEIDIVICEVSGPDLFAGPTCPSPVPAAGDLTADTKIABEROEVSNNEDNR
ITTCRSRIVAGMEBDOI PRNEAOGVHPISIDTHRSNMSPOAMIGKXIYSIG
LIKRFIFGDANITLOADSSSFVNAFPYTTSTKTLTSTRATQODIYYIYLAFRGSM
RIKVALETQDGTGTPRKTNTFTVRMNSLIDSFNSLITSSSAVTTVLPSTINK
GPSTOVIDPTEGLIEVEVPDYPNISHIIPATVITIDGTPSMEDYDKGSPCLLTFSPR
DISATNHIITASFRAVPGDDFSFMYLLEVPPLVAVARA"

ORIGIN
Query Match 65.4%; Score 123.6; DB 13; Length 9185;
Best Local Similarity 51.1%; Pred. No. 6.5e-20;
Matches 95; Conservative 52; Mismatches 39; Indels 0; Gaps 0;
QY 4 AAGUAGUAGUUCUGUCUUAUACAUAUUUGAGAGUUAUAGAAGAGAGUAGUCUA 63
DB 6031 AAAAAATGATCTCTGTTAAATACAAATTTGAGAGGTTAAATTAATTAACAGTAGTCTA 6090
QY 64 UCUAUAUAUAGUUAACUAUUUAUUAUCUGUUCAGAGUCCUUAUUGGACGCCCAU 123
DB 6091 TTTTGTATTTAGGTACTATTAGCTTACGTTCCAGAGATGCTAGTGCAAGCCCA 6150
QY 124 AAUUAUCCAGACACCCUCUCUGUCUUAUAGUAGUAGUUCUUAUUAAGAAGAAA 183
DB 6151 AATATCCAGAGACCTCTCTGCGGTTTTCAGATTAGTAGTCAAAAACCTAAGAAAT 6210
QY 184 UAACCU 189
DB 6211 TTACTT 6216
RESULT 9
BD173516 190 bp RNA linear PAT 18-FEB-2003
LOCUS Novel tertiary structure having ability to accelerate translation
DEFINITION activity.
ACCESSION BD173516
VERSION BD173516.1 GI:28414847
KEYWORDS MO 02061080-A/6.

SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 190)
AUTHORS Nakashima, N. and Kanamori, Y.
TITLE Novel tertiary structure having ability to accelerate translation activity
JOURNAL Patient: MO 02061080-A 6 08-AUG-2002;
JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF
SERICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE
FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA, YASUSHI KANAMORI
OS Black queen-cell virus
PN MO 02061080-A/6
PD 08-AUG-2002
PR 31-JAN-2001 MO 2001JP000641
PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI
PC C12N15/11, C12N15/86, C12P21/02
CC Novel tertiary structure having ability to accelerate CC
translation activity
FH Key location/Qualifiers
FT source 1..190
location/Qualifiers
1..190
/organism="Black queen-cell virus"
/organism="unidentified"
/mol_type="genomic RNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 24.0%; Score 45.4; DB 6; Length 190;
Best Local Similarity 37.7%; Pred. No. 1;
Matches 69; Conservative 41; Mismatches 71; Indels 2; Gaps 2;
QY 7 AUGUAGUUCUGUCUUAUAC-AAUUUGAGAGUUAUAGAAGAGUAGUCUUAUC 65
DB 7 ATGATGATCTCTCTGCGAGGCAAAATTTGACAGATTAATCTGCAAGTAGTGCTATT 66
QY 66 UUAUAUAGUUAACUAUUUAUUAUCUGUUCAGAGUCCCAU-UGGACGCCCAU 124
DB 67 GTTGAATTCACCGTACTATTATTAGATTAGCTCCAGATCGGTGAGTAGCGCTATTC 126
QY 125 AUAUCCAGACACCCUCUCUGUCUUAUUAUAGUAGUAGUUCUUAUUAAGAAGAAU 184
DB 127 AATATCTAGAGAACTGTGCTATTTAGAGATTAGTAGTCTCTAACAAGAACATTT 186
QY 185 AAC 187
DB 187 ACC 189
RESULT 10
BD177020 190 bp RNA linear PAT 16-APR-2003
LOCUS Novel translational activity-promoting higher-order structure.
DEFINITION activity-promoting higher-order structure.
ACCESSION BD177020
VERSION BD177020.1 GI:30014280
KEYWORDS JP 2002306168-A/6.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 190)
AUTHORS Nakashima, N. and Kanamori, Y.
TITLE Novel translational activity-promoting higher-order structure
JOURNAL Patient: JP 2002306168-A 6 22-OCT-2002;
DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND
ENTOMOLOGICAL SCIENCES MINISTRY OF AGRICULTURE FORESTRY AND
ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND
FISHERIES
OS Black queen-cell virus
PN JP 2002306168-A/6
PD 22-OCT-2002
PF 25-JAN-2001 JP 2001016746

FEATURES	source	Location/Qualifiers
ORIGIN		1..190 /organism='Black queen-cell virus'. /organism="unidentified" /mol_type="genomic RNA" /db_xref="taxon:32644"
Query Match	24.0%;	Score 45.4; DB 6; Length 190;
Best Local Similarity	37.7%;	Pred. No. 1;
Matches	69;	Conservative 41; Mismatches 71; Indels 2; Gaps 2
Qy	7	AVUGAUCUUCGUCUUAAC-AUUVUUGAAGGUAUAAGAAGAGUAGUCUAUC 65
Db	7	ANGTCATCTTGGCTTGGGAGGCAAAATTGGACAGATTAATTCGAAATGTCCTATT 66
Qy	66	UUAUAUAGUUAACUAUUAUGUUUAUCUGUUCAGAUGCCUAU-UGCAGCCCCUA 124
Db	67	GTTGAATCAACGATCACTATTAGGTTTACGCTCCAAAGTCGTCGATGACGCCCTATC 126
Qy	125	AUAUCCAGACACCCUUCGUCUUAUUGAUAUAGGUGUCAUUVUUAUAAGAUAU 184
Db	127	AATATCTGAGAAACCTGCTATGTTTAGAAGATTAGTAGTCTTAACAGACAAATT 186
Qy	185	AAC 187
Db	187	ACC 189
RESULT 11		
LOCUS	AF183905	8550 bp ss-RNA linear VRL 16-AUG-2000
DEFINITION		Black queen cell virus nonstructural polyprotein (orf1) and structural polyprotein (orf2) genes, complete cds.
ACCESSION	AF183905	
VERSION	AF183905.1	GI:8100530
KEYWORDS		
SOURCE		Black queen cell virus
ORGANISM		Black queen cell virus
REFERENCE		1 (bases 1 to 8550)
AUTHORS		Leat,N., Ball,B., Govan,V. and Davison,S.
TITLE		Analysis of the complete genome sequence of black queen-cell virus, a picorna-like virus of honey bees
JOURNAL		J. Gen. Virol. 81 (Pt 8), 2111-2119 (2000)
REFERENCE		2 (bases 1 to 8550)
AUTHORS		Leat,N. and Davison,S.
TITLE		Direct Submission
JOURNAL		Submitted (06-SEP-1999) Microbiology, University of Western Cape, Modderdam Rd., Cape Town, Western Cape 7535, South Africa
FEATURES		Location/Qualifiers
source		1..8550 /organism="Black queen cell virus" /mol_type="genomic RNA" /strain="South African" /specific_host="Apis mellifera" /db_xref="taxon:92395" 658..5625 /gene="Orf1" 658..5625 /gene="Orf1" /codon_start=1 /product="nonstructural polyprotein" /protein_id="AAF2337.1" /db_xref="GI:8100531"
CDS		
gene		

[illegible]

Db 5773 AATATCTAGAGAACTGTCATATGTTTAAAGATTAGTACTCTCTAAGACAAATT 5832
QY 185 AAC 187
Db 5833 ACC 5835

RESULT 12
LOCUS CVU55052/c 1347 bp DNA linear VRL 29-JAN-1999
DEFINITION Cowpox virus soluble TNF receptor Crmc (crmc) gene, complete cds.
ACCESSION U55052
VERSION U55052.1 GI:4097321

KEYWORDS
SOURCE Cowpox virus
ORGANISM Cowpox virus
VIRUSES; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae; Orthopoxvirus.

REFERENCE 1 (bases 1 to 1347)
AUTHORS Smith,C.A., Goodwin,R.G. and Pickup,D.J.
TITLE Cowpox virus encodes a second soluble TNF receptor (Crmc) distinct from Crmc

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1347)
AUTHORS Smith,C.A.
TITLE Direct Submission
JOURNAL Submitted (15-APR-1996) Craig A. Smith, Mol. Bio. & Biochem., Immunex, 51 University St., Seattle, WA 98101, USA
LOCATION/Qualifiers

FEATURES
source 1..1347
/organism="Cowpox virus"
/mol_type="genomic DNA"
/strain="Brigton Redgene"
/db_xref="taxon:10243"
357..917
/gene="crmc"
357..917
/gene="crmc"
/note="nonessential gene"
/codon_start=1
/product="soluble TNF receptor Crmc"
/protein_id="AA010325.1"
/db_xref="GI:4097322"

ORIGIN
Query Match 21.4%; Score 40.4; DB 13; Length 1347;
Best Local Similarity 30.9%; Pred. No. 11;
Matches 55; Conservative 37; Mismatches 86; Indels 0; Gaps 0;

QY 10 UGAGUCUCUCUUAUAGUUUGAGGUAUAAGAAGAGUUGUCUAUUA 69
Db 1130 TGAATATATCTTTTAAATTTGAACAAGCTGCTTAAGAGATGTATTGTATTAAAT 1071
QY 70 UAAUUGAGUUAUUAUUGUUUUAUCUGUUCAGAUUGCCUAUUGGACGCCCAUAUAUUC 129
Db 1070 TCATCTGTTTAAATTTATTTCAATAGTATTTTAAATTTTAAATTTTAAATTTTAAAT 1011
QY 130 CAGAGACACCCUUCUUCUUAUUGAUUGAGUUGUCUAUUAAGAUAAGAAUAUAC 187
Db 1010 CAAGAAATTAATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 953

RESULT 13
LOCUS AF482758/c 224499 bp DNA linear VRL 20-MAY-2003
DEFINITION Cowpox virus strain Brigton Red, complete genome.
ACCESSION AF482758 J02066 L08906
VERSION AF482758.2 GI:30795158

KEYWORDS
SOURCE Cowpox virus
ORGANISM Cowpox virus
VIRUSES; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae; Orthopoxvirus.

REFERENCE 1 (bases 1 to 224499)
AUTHORS Pickup,D.J., Bastia,D., Stone,H.O. and Joklik,W.K.
TITLE Sequence of terminal regions of cowpox virus DNA: arrangement of repeated and unique sequence elements
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 79 (23), 7112-7116 (1982)
6961398

JOURNAL 2 (bases 1 to 224499)
AUTHORS Parsons,B.L. and Pickup,D.J.
TITLE Transcription of orthopoxvirus telomeres at late times during infection
JOURNAL Virology 175 (1), 69-80 (1990)
2309453

JOURNAL 3 (bases 1 to 224499)
AUTHORS Hu,F.Q. and Pickup,D.J.
TITLE Transcription of the terminal loop region of vaccinia virus DNA is initiated from the telomere sequences directing DNA resolution
JOURNAL Virology 181 (2), 716-720 (1991)
2014645

JOURNAL 4 (bases 1 to 224499)
AUTHORS Hu,F.Q., Smith,C.A. and Pickup,D.J.
TITLE Cowpox virus contains two copies of an early gene encoding a soluble secreted form of the type II TNF receptor
JOURNAL Virology 204 (1), 343-356 (1994)
8091665

JOURNAL 5 (bases 1 to 224499)
AUTHORS Pickup,D.J.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1993) Molecular Genetics and Microbiology, Duke University Medical Center, Box 3020 DUMC, 421 Jones Building, Durham, NC 27710, USA
6 (bases 1 to 224499)
AUTHORS Dietrich,F.S., Ray,C.A., Sharma,D.A., Allen,A. and Pickup,D.J.
TITLE Direct Submission
JOURNAL Submitted (11-FEB-2002) Molecular Genetics and Microbiology, Duke University Medical Center, Box 3020 DUMC, 421 Jones Building, Durham, NC 27710, USA
7 (bases 1 to 224499)
AUTHORS Dietrich,F.S., Ray,C.A., Sharma,D.A., Allen,A. and Pickup,D.J.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-2003) Molecular Genetics and Microbiology, Duke University Medical Center, Box 3020 DUMC, 421 Jones Building, Durham, NC 27710, USA
On May 16, 2003 this sequence version replaced gi:20152989.

COMMENT
FEATURES
source 1..224499
/organism="Cowpox virus"
/mol_type="genomic DNA"
/strain="Brigton Red"
/db_xref="taxon:10243"
complement(1156..350)
/gene="CPXV001 CDS"
complement(1156..350)
/gene="CPXV001 CDS"
/codon_start=1
/evidence=not experimental
/product="CPXV001 protein"
/protein_id="AA013451.1"
/db_xref="GI:20152989"
/translation="MESISKYYSLSLSFNGLTITLNLFLMESIKYFYSLSLSENGLTKILNLFMVISIKRSIFLTL"
complement(2517..2714)
/gene="CPXV002 CDS"
complement(2517..2714)
/gene="CPXV002 CDS"
/codon_start=1
/evidence=not experimental
/product="CPXV002 protein"
/protein_id="AAP41050.1"

LOCUS AC122491 175092 bp DNA linear ROD 15-MAY-2004
DEFINITION Mus musculus BAC clone RP24-390D3 from chromosome 3, complete
sequence.
ACCESSION AC122491
VERSION AC122491.3 GI:46559826
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 175092)
AUTHORS Nguyen, C., Cotton, M., Bielicki, L. and Meyer, R.
TITLE The sequence of Mus musculus BAC clone RP24-390D3
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 175092)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 175092)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 175092)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 175092)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT On Apr 25, 2004 this sequence version replaced gi:22476248.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.wustl.edu
----- Summary Statistics
Center project name: M_B80390D03

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone, fosmid clone or direct clone walk sequence.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to obtain the consensus sequence; and
the assembly has been confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RPI-24 BAC library has been constructed by Pieter de Jong and
coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is
overlapped by AC138642 and AC111104.
Location/Qualifiers

source

1. .175092
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/map="3"
/clone="RP24-390D3"
/clone_1kb="RPI-24"
251. .977
/rpt_family="ERVX"
repeat_region
1206. .1932
/rpt_family="ERVY"
repeat_region
1972. .3346
/rpt_family="ERVY"
repeat_region
3342. .3857
/rpt_family="ERVY"
repeat_region
3848. .4058
/rpt_family="ERVY"
repeat_region
4059. .4264
/rpt_family="ERVY"
repeat_region
4319. .4598
/rpt_family="ERVY"
repeat_region
6594. .7201
/rpt_family="L1"
repeat_region
7517. .7653
/rpt_family="B4"
repeat_region
8484. .8740
/rpt_family="L1"
repeat_region
8863. .8911
/rpt_family="ERVY"
repeat_region
8951. .9127
/rpt_family="B4"
repeat_region
10944. .11044
/rpt_family="ERVX"
repeat_region
12291. .12631
/rpt_family="L1"
repeat_region
12904. .13112
/rpt_family="ERVX"
repeat_region
13212. .13393
/rpt_family="L1"
repeat_region
13587. .13714
/rpt_family="L1"
repeat_region
17595. .18159
/rpt_family="L1"
repeat_region
18328. .18715
/rpt_family="MALR"
repeat_region
20267. .20423
/rpt_family="A1u"
repeat_region
21854. .22201
/rpt_family="ERVX"
repeat_region
22446. .22569
/rpt_family="L1"
repeat_region
22593. .22941
/rpt_family="L1"
repeat_region
22954. .23328
/rpt_family="L1"
repeat_region
23329. .23478
/rpt_family="A1u"
repeat_region
23479. .23752
/rpt_family="L1"
repeat_region
23793. .24250
/rpt_family="L1"
repeat_region
24272. .24438
/rpt_family="B2"
repeat_region
24565. .24714
/rpt_family="L1"
repeat_region
24746. .24826
/rpt_family="MALR"
repeat_region
24805. .24877
/rpt_family="MALR"
repeat_region
24895. .25240
/rpt_family="MALR"
repeat_region
25273. .28152

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 04:43:02 ; Search time 273.129 seconds
(without alignments)
4611.840 Million cell updates/sec

Title: US-10-088-750C-3

Perfect score: 189
Sequence: 1 guagaagugauguucuu.....uuagaauagaauaacuu 189

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_21:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001s:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189	100.0	189	ABZ80709	Drosophila
2	189	100.0	189	ABZ80709	Drosophila
3	123.6	65.4	188	ABZ80710	Crickel p
4	123.6	65.4	200	ABZ80712	Crickel p
5	45.4	24.0	190	ABZ80712	Black que
6	45.4	24.0	202	ABZ80712	Black que
7	39.8	21.1	43938	AAK77216	Human imm
8	39.8	21.1	45017	AAK77217	Human imm
9	39.8	21.1	64476	ADQ97028	Human can
10	39.6	21.0	5059	ABZ80709	Drosophila
11	39.4	20.8	460	ACH45871	Human foe
12	37.6	19.9	9829	ABZ80711	Aphid let
13	36.4	19.3	186	ABZ80711	Triatoma
14	36.4	19.3	198	ABZ80711	Triatoma
15	36.4	19.0	187	ABZ80708	Himelcobi
16	36.4	19.0	199	ABZ80708	Himelcobi
17	35	18.5	271990	ADQ25213	Fertilily
18	35	18.5	271990	ADQ25213	Fertilily
19	34.8	18.4	37184	ABQ67078	Human ang

c	20	34.6	18.3	4328	2	AAQ28758	AaQ28758 Partial s
c	21	34.4	18.2	130263	6	ABK83573	ABK83573 Human CDN
c	22	34	18.0	6815	10	ADBS4323	ADBS4323 Pretreat
c	23	34	18.0	6815	10	ADBS4207	ADBS4207 Human lym
c	24	34	18.0	12409	6	AAS63312	AAS63312 Chemical1
c	25	33.6	17.8	5273	6	ABL32875	ABL32875 Human imm
c	26	33.4	17.7	5544	6	ABL34620	ABL34620 Human met
c	27	33.4	17.7	5544	6	ABL70477	ABL70477 Chemical1
c	28	33.4	17.7	5544	6	AAS61440	AAS61440 Human gen
c	29	33.4	17.7	5544	7	ADBS9881	ADBS9881 Bialuphic
c	30	33	17.5	18218	6	ABL33949	ABL33949 Human imm
c	31	33	17.5	127432	12	ADQ43653	ADQ43653 Nucleotid
c	32	33	17.5	208765	12	ADQ97430	ADQ97430 Mouse can
c	33	32.8	17.4	6478	10	ADBS4284	ADBS4284 Pretreat
c	34	32.8	17.4	6478	13	ADBS9582	ADBS9582 Oligonuc
c	35	32.6	17.2	17389	6	ABL33415	ABL33415 Human imm
c	36	32.4	17.1	6815	4	AAS45344	AAS45344 Chemical1
c	37	32.4	17.1	6815	6	ABL32670	ABL32670 Human imm
c	38	32.4	17.1	6815	6	ABK28175	ABK28175 DNA trans
c	39	32.4	17.1	6815	10	ADBS4195	ADBS4195 Pretreat
c	40	32.4	17.1	6815	10	ADBS4131	ADBS4131 Human che
c	41	32.4	17.1	7355	6	ABN80297	ABN80297 Human che
c	42	32.4	17.1	9642	6	ABL33357	ABL33357 Human imm
c	43	32.4	17.1	9760	6	ABK31243	ABK31243 Signal tr
c	44	32.4	17.1	9760	6	ABL70198	ABL70198 Chemical1
c	45	32.4	17.1	9760	6	AAS61156	AAS61156 Human gen

ALIGNMENTS

RESULT 1	ABZ80709	standard; RNA, 189 BP.
ID	ABZ80709	standard; RNA, 189 BP.
AC	ABZ80709;	
XX	15-OCT-2003	(first entry)
DT	15-OCT-2003	(first entry)
DE	Drosophila C virus derived pseudoknot sequence.	
XX	Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;	
KW	albumen; impurity; higher-order structure; intergenic region; IGR-IRBS;	
KW	internal ribosome entry site; ss.	
XX	Drosophila C virus.	
OS	Drosophila C virus.	
XX	MO2003033719-A1.	
PN	24-APR-2003.	
PD	08-OCT-2002; 2002MO-JP010447.	
XX	17-OCT-2001; 2001JP-00319923.	
PR	17-OCT-2001; 2001JP-00319923.	
XX	(NAG-) NAT INST AGROBIOLOGICAL SCI.	
PA	(WAKE-) WAKENTYAKU KK.	
XX	Nakashima N, Shibuya N, Nishikawa S;	
PI	WPI; 2003-403230/38.	
DR	Cell-free protein synthesis means in wheatgerm system to establish	
XX	overexpression of target gene with base sequence sustaining translation	
PT	activity and function promotion, for producing useful proteins.	
FT	Claim 1; Page 32; 39pp; Japanese.	
PS	The invention relates to a cell-free protein synthesis system derived	
CC	from wheatgerm where there is substantial exclusion of wheatgerm embryo	
CC	albumen impurities. The novel system uses a sequence having a higher-	
CC	order RNA structure that promotes translation activity. The higher-order	
CC	sequence is preferably a "pseudoknot", especially derived from a range of	

CC viruses (AB280707-AB280713). This sequence represents the "pseudoknot"
 CC higher-order sequence from the Drosophila C virus. The sequence is used
 CC in a construct which may also include an intergenic region and internal
 CC ribosome entry site (IGR-IRRS). The method is applicable in producing
 CC useful proteins

CC Sequence 189 BP; 57 A; 30 C; 34 G; 0 T; 68 U; 0 Other;

Query Match 100.0%; Score 189; DB 8; Length 189;
 Best Local Similarity 100.0%; Pred. No. 2.4e-43;
 Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUUAGAUGUGAUCUGUCUCCUUAUACAUUUUGAGGUAUAAGAAGAAAGUAGUG 60
 DB 1 GUUAGAUGUGAUCUGUCUCCUUAUACAUUUUGAGGUAUAAGAAGAAAGUAGUG 60
 QY 61 CUUUCUUAUAUAUUGAGUUAACUUAUUUGUUUAUCUGUUCAGAGUCCUUAUUGGACGCC 120
 DB 61 CUUUCUUAUAUAUUGAGUUAACUUAUUUGUUUAUCUGUUCAGAGUCCUUAUUGGACGCC 120
 QY 121 CAUAAUUAUCCAGACACCCUCUCUCUCUUAUUAUUAUUAUUAUUAUUAUUAUUA 180
 DB 121 CAUAAUUAUCCAGACACCCUCUCUCUCUUAUUAUUAUUAUUAUUAUUAUUAUUA 180
 QY 181 AAUUAACCU 189
 DB 181 AAUUAACCU 189

RESULT 2
 AAL50546
 ID AAL50546 standard; RNA; 201 BP.

AC AAL50546;
 XX
 XX 19-DEC-2002 (first entry)
 XX
 XX Drosophila C virus RNA sequence.

XX CrPV-like virus; ss; higher-order structure; drug development;
 KM drug production; translational activity-promoting function;
 KM protein synthesis; structural analysis.

XX Drosophila C virus.

XX WO200261080-A1.

XX 08-AUG-2002.

XX 31-JAN-2001; 2001WO-JP000641.

XX 25-JAN-2001; 2001JP-00016746.

XX (NAG-) NAT INST AGROBIOLOGICAL SCI.

XX Nakashima N, Kanamori Y;

XX WPI; 2002-627482/67.

XX Translational activity-promoting higher-order structure of CrPV-like
 PT viruses for protein translation when suitably inlaid, useful in
 PT synthesis of proteins and polypeptides of foreign species for application
 in drugs.

XX Claim 1; Fig 1-2; 38pp; Japanese.

XX The invention comprises seven RNA sequences (CrPV-like viruses) which
 CC have a higher-order structure that sustains translational activity-
 CC promoting function. The RNA sequences of the invention are useful in the
 CC synthesis of proteins and polypeptides for application in developing and
 CC producing drugs. The RNA sequences of the invention are also useful in
 CC basic research of protein synthesis and structural analysis by the gene
 CC recombinant technique. The present nucleotide represents a Drosophila C

CC virus RNA sequence of the invention

XX Sequence 201 BP; 61 A; 33 C; 35 G; 0 T; 72 U; 0 Other;

Query Match 100.0%; Score 189; DB 6; Length 201;
 Best Local Similarity 100.0%; Pred. No. 2.5e-43;
 Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUUAGAUGUGAUCUGUCUCCUUAUACAUUUUGAGGUAUAAGAAGAAAGUAGUG 60
 DB 1 GUUAGAUGUGAUCUGUCUCCUUAUACAUUUUGAGGUAUAAGAAGAAAGUAGUG 60
 QY 61 CUUUCUUAUAUAUUGAGUUAACUUAUUUGUUUAUCUGUUCAGAGUCCUUAUUGGACGCC 120
 DB 61 CUUUCUUAUAUAUUGAGUUAACUUAUUUGUUUAUCUGUUCAGAGUCCUUAUUGGACGCC 120
 QY 121 CAUAAUUAUCCAGACACCCUCUCUCUCUUAUUAUUAUUAUUAUUAUUAUUAUUA 180
 DB 121 CAUAAUUAUCCAGACACCCUCUCUCUCUUAUUAUUAUUAUUAUUAUUAUUAUUA 180
 QY 181 AAUUAACCU 189
 DB 181 AAUUAACCU 189

RESULT 3
 AB280710
 ID AB280710 standard; RNA; 188 BP.

AC AB280710;
 XX
 XX 15-OCT-2003 (first entry)
 XX
 XX Cricket paralysis virus derived pseudoknot sequence.

XX Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;
 KM albumen; impurity; higher-order structure; intergenic region; IGR-IRRS;
 KM internal ribosome entry site; ss.

XX Cricket paralysis virus.

XX WO2003033719-A1.

XX 24-APR-2003.

XX 08-OCT-2002; 2002WO-JP010447.

XX 17-OCT-2001; 2001JP-00319923.

XX (NAG-) NAT INST AGROBIOLOGICAL SCI.

XX (WAKE-) WAKENYAKU KK.

XX Nakashima N, Shibuya N, Nishikawa S;

XX WPI; 2003-403230/38.

XX Cell-free protein synthesis means in wheatgerm system to establish
 PT overexpression of target gene with base sequence sustaining translation
 PT activity and function promotion, for producing useful proteins.

XX Claim 1; Page 33; 39pp; Japanese.

XX The invention relates to a cell-free protein synthesis system derived
 CC from wheatgerm where there is substantial exclusion of wheatgerm embryo
 CC albumen impurities. The novel system uses a sequence having a higher-
 CC order RNA structure that promotes translation activity. The higher-order
 CC sequence is preferably a "pseudoknot", especially derived from a range of
 CC viruses (AB280707-AB280713). This sequence represents the "pseudoknot"
 CC higher-order sequence from the Cricket paralysis virus. The sequence is
 CC used in a construct which may also include an intergenic region and
 CC internal ribosome entry site (IGR-IRRS). The method is applicable in
 CC producing useful proteins

Sequence 188 BP; 56 A; 33 C; 36 G; 0 T; 63 U; 0 Other;
Query Match 65.4%; Score 123.6; DB 8; Length 188;
Best Local Similarity 79.0%; Fred. No. 5.8e-25;
Matches 147; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
Db 4 AAGUUGAUCUUCGUCUUAUACAUAUUUGAGAGUUAAUAGAAGAGUAGUGCUA 63
3 AAUUGUAUCUUCGUCUUAUACAUAUUUGAGAGUUAAUAGAAGAGUAGUGCUA 62
Qy 64 UCUUAAUUAUAGGUUAACUAUUUUAUGUUUAUGUUGCAGAUUGCCUUAUUGCCAGCCCAU 123
Db 63 UUUUUGUUAUUUAGGUUAGCUUUUUUAGCUUUAUGUUGCAGAUUGCCUUAUUGCCAGCCCAU 122
Qy 124 AAUUAUCCAGACACCCUUCUGUCUUAUUAUGAUUAGUGUUCUUAUUAAGAAUAA 183
Db 123 AAUUAUCCAGAGAGCCUUCUGCGGUUUUUUUCAGAUUAGUAGUAGCAAAAACCUAAGAAU 182
Qy 184 UAACCU 189
Db 183 UUAACCU 188
RESULT 4
AAL50547
ID AAL50547 standard; RNA; 200 BP.
AC AAL50547;
DT 19-DEC-2002 (first entry)
DE Cricket paralysis virus RNA sequence.
XX CrPV-like virus; ss; higher-order structure; drug development;
KM drug production; translational activity-promoting function;
KM protein synthesis; structural analysis.
XX Cricket paralysis virus.
OS
XX WO200261080-A1.
PN 08-AUG-2002.
PD 31-JAN-2001; 2001WO-JP000641.
PR 25-JAN-2001; 2001JP-00016746.
XX (NAG-) NAT INST AGROBIOLOGICAL SCI.
PA Nakashima N, Kanamori Y;
PI WPI; 2002-627482/67.
DR Translational activity-promoting higher-order structure of CrPV-like
XX PT viruses for protein translation when suitably initiated, useful in
PT synthesis of proteins and polypeptides of foreign species for application
PT in drugs.
XX Claim 1; Fig 1-2; 38pp; Japanese.
PS The invention comprises seven RNA sequences (CrPV-like viruses) which
XX have a higher-order structure that sustains translational activity-
CC promoting function. The RNA sequences of the invention are useful in the
CC synthesis of proteins and polypeptides for application in developing and
CC producing drugs. The RNA sequences of the invention are also useful in
CC basic research of protein synthesis and structural analysis by the gene
CC recombinant technique. The present nucleotide represents a Cricket
CC paralysis virus RNA sequence of the invention
XX Sequence 200 BP; 60 A; 36 C; 37 G; 0 T; 67 U; 0 Other;
Query Match 65.4%; Score 123.6; DB 8; Length 200;
Best Local Similarity 79.0%; Fred. No. 5.8e-25;

Matches 147; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
Qy 4 AAGUUGAUCUUCGUCUUAUACAUAUUUGAGAGUUAAUAGAAGAGUAGUGCUA 63
Db 3 AAUUGUAUCUUCGUCUUAUACAUAUUUGAGAGUUAAUAGAAGAGUAGUGCUA 62
Qy 64 UCUUAAUUAUAGGUUAACUAUUUUAUGUUUAUGUUGCAGAUUGCCUUAUUGCCAGCCCAU 123
Db 63 UUUUUGUUAUUUAGGUUAGCUUUUUUAGCUUUAUGUUGCAGAUUGCCUUAUUGCCAGCCCAU 122
Qy 124 AAUUAUCCAGACACCCUUCUGUCUUAUUAUGAUUAGUGUUCUUAUUAAGAAUAA 183
Db 123 AAUUAUCCAGAGAGCCUUCUGCGGUUUUUUUCAGAUUAGUAGUAGCAAAAACCUAAGAAU 182
Qy 184 UAACCU 189
Db 183 UUAACCU 188
RESULT 5
ABZ80712
ID ABZ80712 standard; RNA; 190 BP.
AC ABZ80712;
DT 15-OCT-2003 (first entry)
DE Black queen-cell virus derived pseudoknot sequence.
XX Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;
KM albumen; impurity; higher-order structure; intergenic region; IGR-IRES;
KM internal ribosome entry site; ss.
XX Black queen-cell virus.
OS
XX WO2003033719-A1.
PN 24-APR-2003.
PD 08-OCT-2002; 2002WO-JP010447.
PR 17-OCT-2001; 2001JP-00319923.
XX (NAG-) NAT INST AGROBIOLOGICAL SCI.
PA (WAKE-) WAKENYAKU KK.
PI Nakashima N, Shibuya N, Nishikawa S;
DR WPI; 2003-403230/38.
XX Cell-free protein synthesis means in wheatgerm system to establish
PT overexpression of target gene with base sequence sustaining translation
PT activity and function promotion, for producing useful proteins.
XX Claim 1; Page 34; 39pp; Japanese.
PS The invention relates to to a cell-free protein synthesis system derived
XX from wheatgerm where there is substantial exclusion of wheatgerm embryo
CC albumen impurities. The novel system uses a sequence having a higher-
CC order RNA structure that promotes translation activity. The higher-order
CC sequence is preferably a "pseudoknot", especially derived from a range of
CC viruses (ABZ80707-ABZ80713). This sequence represents the "pseudoknot"
CC higher-order sequence from the Black queen-cell virus. The sequence is
CC used in a construct which may also include an intergenic region and
CC internal ribosome entry site (IGR-IRES). The method is applicable in
CC producing useful proteins
XX Sequence 190 BP; 58 A; 36 C; 40 G; 0 T; 56 U; 0 Other;
Query Match 24.0%; Score 45.4; DB 8; Length 190;
Best Local Similarity 60.1%; Fred. No. 0.0054;
Matches 110; Conservative 0; Mismatches 71; Indels 2; Gaps 2;


```
QY 7 AUGGUAUCUUGUCUUCUUAUAC-AAUUTUGAGAGUUAUAGAAGAGUAGUCUUAUC 65
DB 7 AUGGUAUCUUGUCUUCGCGAGCGCAAAAUUUGCAACAUAUAAAUCUGCAAGAGUCUUAU 66
QY 66 UUAUAUUAUAGGUAUACUUAUUUAAGUUUACUGUUCAGAGUCCUUAU-UGGCAGCCCCAUA 124
DB 67 GUGGAAUUCACCGUACCUUAUUUAAGUUUACGUCUCCAGAGUCCGUGAGUAGCAGCCCUUAUC 126
QY 125 AUUUCAGAGACACCCUUCUCUCUUCUUAUUGAUUUAAGUUGUCUUAUUAAGAAUUA 184
DB 127 AAUUAUCUAGAGAACUUGUCUUAUGUUUAGAAUUAAGUAGUCUCUUAACAGAACAAU 186
QY 185 AAC 187
DB 187 ACC 189

RESULT 6
ID AAL50549 standard; RNA; 202 BP.
XX AAL50549;
AC AAL50549;
DT 07-AUG-2003 (revised)
DT 19-DEC-2002 (first entry)
XX
XX Black queen-cell virus RNA sequence.
DE
XX CrPV-like virus; ss; higher-order structure; drug development;
KM drug production; translational activity-promoting function;
XX protein synthesis; structural analysis.
XX
XX Black queen cell virus.
OS
XX WO200261080-A1.
XX
XX 08-AUG-2002.
PD
XX 31-JAN-2001; 2001WO-JP000641.
PF
XX 25-JAN-2001; 2001JP-00016746.
PR
XX (NNAAG-) NAT INST AGRBIOLOGICAL SCI.
XX
XX Nakashima N, Kanamori Y;
PI
XX WPI; 2002-627482/67.
DR
XX
XX Translational activity-promoting higher-order structure of CrPV-like
PT viruses for protein translation when suitably initiated, useful in
PT synthesis of proteins and polypeptides of foreign species for application
PT in drugs.
XX
XX Claim 1; Fig 1-2; 38pp; Japanese.
XX
XX The invention comprises seven RNA sequences (CrPV-like viruses) which
CC have a higher-order structure that sustains translational activity-
CC promoting function. The RNA sequences of the invention are useful in the
CC synthesis of proteins and polypeptides for application in developing and
CC producing drugs. The RNA sequences of the invention are also useful in
CC basic research of protein synthesis and structural analysis by the gene
CC recombinant technique. The present nucleotide sequence represents a Black queen-
CC cell virus RNA sequence of the invention. (Updated on 07-AUG-2003 to
CC correct OS field.)
XX
XX
SQ Sequence 202 BP; 63 A; 38 C; 42 G; 0 T; 59 U; 0 Other;
Query Match 24.0%; Score 45.4; DB 6; Length 202;
Best Local Similarity 60.1%; Pred. No. 0.0055;
Matches 110; Conservative 0; Mismatches 71; Indels 2; Gaps 2;
```

```
DB 7 AUGGUAUCUUGUCUUCGCGAGCGCAAAAUUUGCAACAUAUAAAUCUGCAAGAGUCUUAU 66
QY 66 UUAUAUUAUAGGUAUACUUAUUUAAGUUUACUGUUCAGAGUCCUUAU-UGGCAGCCCCAUA 124
DB 67 GUGGAAUUCACCGUACCUUAUUUAAGUUUACGUCUCCAGAGUCCGUGAGUAGCAGCCCUUAUC 126
QY 125 AUUUCAGAGACACCCUUCUCUCUUCUUAUUGAUUUAAGUUGUCUUAUUAAGAAUUA 184
DB 127 AAUUAUCUAGAGAACUUGUCUUAUGUUUAGAAUUAAGUAGUCUCUUAACAGAACAAU 186
QY 185 AAC 187
DB 187 ACC 189

RESULT 7
ID AAK77216 standard; DNA; 43938 BP.
XX AAK77216;
AC AAK77216;
DT 07-NOV-2001 (first entry)
DT
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32028.
DE
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KM cytostatic; gene therapy; vaccine; metastasis; de.
XX
XX Homo sapiens.
OS
XX WO200157182-A2.
XX
XX 09-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US0001354.
PF
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216479P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 16-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0229824P.
PR 01-SEP-2000; 2000US-0229287P.
```


	PR	01-SEP-2000;	2000US-02299343P.
	PR	01-SEP-2000;	2000US-02299343P.
	PR	01-SEP-2000;	2000US-02299343P.
	PR	01-SEP-2000;	2000US-02299343P.
	PR	05-SEP-2000;	2000US-0229509P.
	PR	05-SEP-2000;	2000US-0229509P.
	PR	06-SEP-2000;	2000US-0230437P.
	PR	06-SEP-2000;	2000US-0230437P.
	PR	08-SEP-2000;	2000US-0231242P.
	PR	08-SEP-2000;	2000US-0231243P.
	PR	08-SEP-2000;	2000US-0231243P.
	PR	08-SEP-2000;	2000US-0231413P.
	PR	08-SEP-2000;	2000US-0231413P.
	PR	08-SEP-2000;	2000US-0232080P.
	PR	08-SEP-2000;	2000US-0232081P.
	PR	12-SEP-2000;	2000US-0231968P.
	PR	14-SEP-2000;	2000US-0232397P.
	PR	14-SEP-2000;	2000US-0232398P.
	PR	14-SEP-2000;	2000US-0232399P.
	PR	14-SEP-2000;	2000US-0232400P.
	PR	14-SEP-2000;	2000US-0232401P.
	PR	14-SEP-2000;	2000US-0233063P.
	PR	14-SEP-2000;	2000US-0233063P.
	PR	14-SEP-2000;	2000US-0233065P.
	PR	21-SEP-2000;	2000US-0234223P.
	PR	21-SEP-2000;	2000US-0234224P.
	PR	25-SEP-2000;	2000US-0234997P.
	PR	25-SEP-2000;	2000US-0234998P.
	PR	26-SEP-2000;	2000US-0235834P.
	PR	27-SEP-2000;	2000US-0235836P.
	PR	29-SEP-2000;	2000US-0236327P.
	PR	29-SEP-2000;	2000US-0236327P.
	PR	29-SEP-2000;	2000US-0236368P.
	PR	29-SEP-2000;	2000US-0236369P.
	PR	29-SEP-2000;	2000US-0236370P.
	PR	02-OCT-2000;	2000US-0237037P.
	PR	02-OCT-2000;	2000US-0237037P.
	PR	02-OCT-2000;	2000US-0237038P.
	PR	02-OCT-2000;	2000US-0237039P.
	PR	13-OCT-2000;	2000US-0237040P.
	PR	13-OCT-2000;	2000US-0239935P.
	PR	13-OCT-2000;	2000US-0239937P.
	PR	20-OCT-2000;	2000US-0240960P.
	PR	20-OCT-2000;	2000US-0241221P.
	PR	20-OCT-2000;	2000US-0241785P.
	PR	20-OCT-2000;	2000US-0241786P.
	PR	20-OCT-2000;	2000US-0241787P.
	PR	20-OCT-2000;	2000US-0241808P.
	PR	20-OCT-2000;	2000US-0241809P.
	PR	20-OCT-2000;	2000US-0241826P.
	PR	01-NOV-2000;	2000US-0244617P.
	PR	08-NOV-2000;	2000US-0246474P.
	PR	08-NOV-2000;	2000US-0246475P.
	PR	08-NOV-2000;	2000US-0246476P.
	PR	08-NOV-2000;	2000US-0246477P.
	PR	08-NOV-2000;	2000US-0246523P.
	PR	08-NOV-2000;	2000US-0246524P.
	PR	08-NOV-2000;	2000US-0246525P.
	PR	08-NOV-2000;	2000US-0246526P.
	PR	08-NOV-2000;	2000US-0246527P.
	PR	08-NOV-2000;	2000US-0246528P.
	PR	08-NOV-2000;	2000US-0246532P.
	PR	08-NOV-2000;	2000US-0246609P.
	PR	08-NOV-2000;	2000US-0246610P.
	PR	08-NOV-2000;	2000US-0246611P.
	PR	08-NOV-2000;	2000US-0246613P.
	PR	17-NOV-2000;	2000US-0249207P.
	PR	17-NOV-2000;	2000US-0249208P.
	PR	17-NOV-2000;	2000US-0249209P.
	PR	17-NOV-2000;	2000US-0249210P.
	PR	17-NOV-2000;	2000US-0249211P.
	PR	17-NOV-2000;	2000US-0249212P.
	PR	17-NOV-2000;	2000US-0249213P.
	PR	17-NOV-2000;	2000US-0249214P.
	PR	17-NOV-2000;	2000US-0249215P.
	PR	17-NOV-2000;	2000US-0249216P.
	PR	17-NOV-2000;	2000US-0249217P.
	PR	17-NOV-2000;	2000US-0249218P.
	PR	17-NOV-2000;	2000US-0249244P.
	PR	17-NOV-2000;	2000US-0249245P.
	PR	17-NOV-2000;	2000US-0249246P.
	PR	17-NOV-2000;	2000US-0249265P.
	PR	17-NOV-2000;	2000US-0249297P.
	PR	17-NOV-2000;	2000US-0249299P.
	PR	17-NOV-2000;	2000US-0249300P.
	PR	01-DEC-2000;	2000US-0250391P.
	PR		

RESULT 8
AAK7217
ID AAK7217 standard; DNA; 45017 BP.
XX AAK7217;
AC
XX AAK7217;
DT 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32029.
DE
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW
KW Cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
OS
XX W0200157182-A2.
FN
XX
XX 09-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US001354.
PF
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 14-AUG-2000; 2000US-0226279P.
PR 18-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-0239357P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246533P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.

PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX
XX Disclosure; SEQ ID NO 32029; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK67694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 45017 BP; 12919 A; 10037 C; 9566 G; 12495 T; 0 U; 0 Other;
Query Match 21.1%; Score 39.8; DB 4; Length 45017;
Best Local Similarity 32.2%; Pred. No. 0.69;
Matches 55; Conservative 34; Mismatches 82; Indels 0; Gaps 0;
QY 10 UGAUCUCGUCUCCUUAUACAUAUUUGAGAGUUAUAAGAAGAGUAGUUGCUUAUUA 69
DB 8574 TCATCTCGGCTTAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8633
QY 70 UAUUAGAGUUAUAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 129
DB 8634 AAATTAAGTCACATTTTAACTTTTGTGTGAATTTTCTACAGTTCCTCAAAATTTT 8633
QY 130 CAGGACACCCUUCUUCUUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUA 180
DB 8694 CAGTAAGTTCTGCCCCCTTGTCAAGTTATCTGCGACGCAAAAAA 8744
RESULT 9
ADQ97028
XX ADQ97028 standard; DNA; 84476 BP.
XX
XX ADQ97028;
DT 07-OCT-2004 (first entry)
XX
XX Human cancer associated sequence HD10-002, SEQ ID 4.
DE
XX
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.

XX
OS Homo sapiens.
XX
XX WO2004060304-A2.
XX
XX 22-JUL-2004.
XX
XX 22-DEC-2003; 2003WO-US041389.
XX
XX 27-DEC-2002; 2002US-00330773.
XX
XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Malandro MS;
XX
XX WPI; 2004-543781/52.
XX
XX New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.
XX
XX
XX Claim 1; SEQ ID NO 4; 19pp; English.
XX
XX The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic formate directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 84476 BP; 22346 A; 17371 C; 17200 G; 27559 T; 0 U; 0 Other;
Query Match 21.1%; Score 39.8; DB 12; Length 84476;
Best Local Similarity 32.2%; Pred. No. 0.79;
Matches 55; Conservative 34; Mismatches 82; Indels 0; Gaps 0;
QY 10 UGAUCUCGUCUCCUUAUACAUAUUUGAGAGUUAUAAGAAGAGUAGUUGCUUAUUA 69
DB 83402 TCATCTCGGCTTAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 83461
QY 70 UAUUAGAGUUAUAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 129
DB 83462 AAATTAAGTCACATTTTAACTTTTGTGTGAATTTTCTACAGTTCCTCAAAATTTT 83521
QY 130 CAGGACACCCUUCUUCUUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUA 180
DB 83522 CAGTAAGTTCTGCCCCCTTGTCAAGTTATCTGCGACGCAAAAAA 83572
RESULT 10
ABL18538/C
XX ABL18538 standard; DNA; 5059 BP.
XX
XX ABL18538;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 7087.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX

(PLAN-) PLANT RES INT BV.

Van Der Wilk F, Van Munster M;
WPI; 2003-441523/41.

Novel nucleic acid derived from aphid lethal paralysis virus, useful for detecting ALPV-viruses, for preparing a monoclonal antibody against antigen, in research concerning aphids and for controlling aphids.

Claim 1; Page 10-13; 17pp; English.

The invention relates to a novel RNA derived from an ALPV (aphid lethal paralysis virus). ALPV RNA is useful for controlling homopterous insects such as aphids or white flies. It is useful in the research concerning aphids and for controlling aphids and in the field of agriculture or floriculture. It is also useful for detecting ALPV-viruses using PCR and for preparing a monoclonal antibody against antigen which is synthesised in vitro. The invention is also useful in gene therapy. The present sequence is ALPV RNA

Sequence 9829 BP; 3083 A; 1906 C; 1879 G; 0 T; 2961 U; 0 Other;

Query Match 19.9%; Score 37.6; DB 8; Length 9829;
Best Local Similarity 62.5%; Pred. No. 2;
Matches 75; Conservative 0; Mismatches 44; Indels 1; Gaps 1;

62 UAUCUUAUUAAUUGAGUUAACUAVUUGUUUACUGUUCAGAUUGCAGGCCCC 121
UUUGAUAUUUUAAGGCUUAGUUAUUUAACUUAACUUAACAAGAGCGCCGUGCAGCCCC 6756

122 AUAAUUAUCCAGACAGACCUCUCUCGUCCUUAUUAUGAUUGAGUUGUCAUUUAGAUAAGA 181
ACGAUAUCCUAAUGUUAAGUCCGAUG-UCCUAUUUUGAUUAGUGUGUCAUUGGUCAGAAAA 6815

RESULT 13

ABZ80711
ID ABZ80711 standard; RNA; 186 BP.
XX ABZ80711;
DT 15-OCT-2003 (first entry)
DE Triatoma virus derived pseudoknot sequence.
KW Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm; albumen; impurity; higher-order structure; intergenic region; IGR-IRBS;
XX internal ribosome entry site; ss.
OS Triatoma virus.
PN WO2003033719-A1.
PD 24-APR-2003.
PF 08-OCT-2002; 2002WO-JP010447.
PR 17-OCT-2001; 2001JP-00319923.
PA ((NAA-G)) NAT INST AGROBIOLOGICAL SCI.
PA ((WAKE-)) WAKENYAKU KK.
PI Nakashima N, Shibuya N, Nishikawa S;
DR WPI; 2003-403230/38.
XX Cell-free protein synthesis means in wheatgerm system to establish overexpression of target gene with base sequence sustaining translation activity and function promotion, for producing useful proteins.
PS Claim 1; Page 33; 39pp; Japanese.

CC	The invention relates to a cell-free protein synthesis system derived
CC	from wheatgerm where there is substantial exclusion of wheatgerm cyto
CC	.albumen impurities. The novel system uses a sequence having a higher-
CC	order RNA structure that promotes translation activity. The higher-order
CC	sequence is preferably a "pseudoknot", especially derived from a range of
CC	viruses (AB880707-AB880713). This sequence represents the "pseudoknot"
CC	higher-order sequence from the Triatoma virus. The sequence is used in a
CC	construct which may also include an intergenic region and internal
CC	ribosome entry site (IGR-IRES). The method is applicable in producing
CC	useful proteins
SQ	Sequence 186 BP; 49 A; 35 C; 42 G; 0 T; 60 U; 0 Other;
DY	Query Match 19.3%; Score 36.4; DB 8; Length 186;
DY	Best Local Similarity 59.8%; Pred. No. 1.8;
DY	Matches 79; Conservative 0; Mismatches 51; Indels 2; Gaps 1
DY	2 UUAAGUUGGAGUCUUGCGUCCUUACAAUUUUUGAGAGUUAAUAGAAGAGUAGUC 61
DY	2 UGACUUGUGAUCUGCUCUUCCGAUAAAUCUGACUAAAAGUCGAAAUUUGCU-- 59
DY	62 UAUUUUAUUUAUUGGUUAACUAAUUUUAUUGUUGCAAGUUGCCAUUUGGAGCCCC 121
DY	60 AUAGUUAAAGUUGGCCGUGCCAUUUUAGGCADVACUUCUCAGAGUGGGCGUGCAGCCA 119
DY	122 AUAUAUCCAGG 133
DY	120 ACAAGAVCCAGG 131
RESULT 14	
AAL50548	
ID	AAL50548 standard; RNA; 198 BP.
XX	
AC	AAL50548;
XX	
DT	19-DEC-2002 (first entry)
XX	
DE	Triatoma virus RNA sequence.
XX	
KM	CrPV-like virus; ss; higher-order structure; drug development;
KM	drug production; translational activity-promoting function;
KW	protein synthesis; structural analysis.
XX	
OS	Triatoma virus.
XX	
PN	WO200261080-A1.
XX	
PD	08-AUG-2002.
XX	
PF	31-JAN-2001; 2001WO-JP000641.
PR	25-JAN-2001; 2001JP-00016746.
PA	(NAAG-) NAT INST AGROBIOLOGICAL SCI.
PI	Nakashima N, Kanamori Y;
PT	WPI; 2002-627482/67.
DR	
XX	
PT	Translational activity-Promoting higher-order structure of CrPV-like
PT	viruses for protein translation when suitably initiated, useful in
PT	synthesis of proteins and polypeptides of foreign species for application
PT	in drugs.
PS	Claim 1, Fig 1-2, 3bpp; Japanese.
CC	The invention comprises seven RNA sequences (CrPV-like viruses) which
CC	have a higher-order structure that sustains translational activity-
CC	promoting function. The RNA sequences of the invention are useful in the
CC	synthesis of proteins and polypeptides for application in developing and
CC	producing drugs. The RNA sequences of the invention are also useful in
CC	basic research of protein synthesis and structural analysis by the gene

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 07:31:33 : Search time 1977.76 seconds
(without alignments)
4471.109 Million cell updates/sec

Title: US-10-088-750C-3

Perfect score: 189
Sequence: 1 guuaagugugauucugcu.....uuagaauagaauaaccu 189

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	187.4	99.2	433	7	CO153155 EN01401.5
2	187.4	99.2	486	7	CO153454 EN02010.5
3	187.4	99.2	490	7	CO153765 EN02756.5
4	187.4	99.2	539	7	CO337794 EN15317.5
5	68	36.0	453	7	CO338064 EN15862.5
6	48	25.4	472	7	CO337705 EN15125.5
7	44.2	23.4	448	7	CO224988 WS01026.B
8	42.2	22.3	873	10	C2956789 274903 TO
9	40	21.2	748	9	A2748753 RPCI-24-6
10	40	21.2	942	10	C2338852 ZMMBFP0107
11	39.4	20.8	1642	10	AG520404 Mus muscu
12	39.2	20.7	351	7	CN253956 BIP018 B
13	39.2	20.7	351	7	CN254287 BIP0433 B
14	38.6	20.4	1088	10	CL045570 CH216-63A
15	38.4	19.8	524	9	AG430181 Mus muscu
16	37.4	19.8	524	9	AQ392257 CITR1-E1-
17	37.4	19.8	714	9	AQ040332 CIT-HSP-2
18	37.2	19.7	545	9	BH32653 CH230-124
19	37	19.6	716	7	CJ404651 CJ404651
20	36.6	19.4	729	7	AQ075710 CIT-HSP-2
21	36.6	19.4	1008	10	CL483414 SATL_37b
22	36.4	19.3	656	10	AG311732 Mus muscu

23	36.4	19.3	842	9	CC070286 CSU-K33r.
24	36.4	19.3	925	10	C2506071 GMR2-13B1
25	36.4	19.3	1057	10	CL997184 ZMMBHC001
26	36.2	19.2	394	2	BF908674 IL2-UT007
27	36.2	19.2	811	9	B2385506 B15-874ST
28	36.2	19.2	814	9	CC972368 ZUAEK42TV
29	36.2	19.2	816	9	CC546802 CH240_429
30	36	19.0	541	9	AQ419640 RPCI-11-1
31	36	19.0	547	6	CD451731 USDA-FP-1
32	36	19.0	563	10	CZ705112 OC_Ba002
33	36	19.0	682	3	BM168656 EST571179
34	36	19.0	684	10	CZ783049 OC_Ba014
35	36	19.0	1039	10	CG418022 ZMMBHC002
36	35.8	18.9	284	3	BF191829 603307551
37	35.8	18.9	978	9	CC717480 OGPB163TV
38	35.6	18.8	769	9	BZ489474 BONFG30TF
39	35.6	18.8	921	10	CL099026 ISB1-32B1
40	35.4	18.7	907	10	DU002474 284709 TO
41	35.2	18.6	675	9	AZ527316 267PB502
42	35.2	18.6	784	6	CB199345 AGENCOURT
43	35.2	18.6	789	7	CK316785 SB02019B2
44	35.2	18.6	975	8	DN562098 90906553
45	35	18.5	364	9	BZ216751 CH230-250

ALIGNMENTS

RESULT 1
CO153155 433 bp mRNA linear EST 05-OCT-2004
LOCUS EN01401.5prine Exelixis FlyTag MN08 Bluescript Drosophila
DEFINITION melanogaster cDNA clone EN01401 5, mRNA sequence.
ACCESSION CO153155
VERSION CO153155.1 GI:48907156
KEYWORDS
SOURCE
ORGANISM Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 433)
Nakanishi, M., Muzong, C., Peterson, E., Laufer, A., Leung, W., Platt, D.
and Swimmer, C.
REFERENCE Exelixis FlyTag EST Project MN08 Library
JOURNAL Unpublished (2004)
COMMENT Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
Plate: EN.14 row: A column: 1
High quality sequence stop: 383.
Location/Qualifiers
1..433
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="EN01401"
/cell_line="mbn2"
/note="Vector: pBluescript, Site 1: NotI, Site 2: XhoI;
oligoRT primed from LPS induced mbn2 cell line."

FEATURES

source

ORIGIN

Query Match 99.2%, Score 187.4; DB 7; Length 433;
Best Local Similarity 63.5%; Pred. No. 1.2e-41;
Matches 120; Conservative 68; Mismatches 1; Indels 0; Gaps 0;

0y 1 GUUAAGUGUGAUCUCGUCUUAUACAUAUUGAGAGUUAUAAGAGAGUAGUG 60
Db 87 GTTAAAGATGATCTTCCTTATACATTGAGAGGTTAATAAGAGAGATGATG 146

[illegible]

LOCUS	CO153765	490 bp	mRNA	linear	EST 05-OCT-2004
DEFINITION	CO153765.5prime Exelixis FlyTag MN08 Bluescript Drosophila melanogaster cDNA clone EN02756 5, mRNA sequence.				
ACCESSION	CO153765				
VERSION	CO153765.1	GI:48907766			
KEYWORDS	EST.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Archopoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
AUTHORS	1 (bases 1 to 490)				
TITLE	Nakanishi, M., Muzong, C., Peterson, E., Laufer, A., Leung, W., Platt, D. and Swamy, C.				
JOURNAL	Exelixis FlyTag EST Project MN08 Library				
COMMENT	Unpublished (2004) Contact: Stapleton, M. BDGP Lawrence Berkeley National Lab One Cyclotron Rd, Berkeley, CA 94720, USA Fax: 510 486 6798 Email: http://www.fruitfly.org/EST , est@fruitfly.berkeley.edu Plate: EN.27 row: E column: 8 High quality sequence stop: 401. Location/Qualifiers 1..490 /organism="Drosophila melanogaster" /mol_type="mRNA" /db_xref="taxon:7227" /clone="EN02756" /cell_line="mbn2" /clone_id="Exelixis FlyTag MN08 Bluescript" /note="Vector: plusscript; Site 1: NotI; Site 2: XhoI; oligodt primed from LPS induced mbn2 cell line."				
FEATURES	source				
ORIGIN					
Query Match	99.2%; Score 187.4; DB 7; Length 490;				
Best Local Similarity	63.5%; Pred. No. 1.2e-41;				
Matches 120; Conservative	68; Mismatches 1; Indels 0; Gaps 0;				
QY	1 GUNAAAGUGAUCUUCUCCUCCUUAUACAUAUUGAGAGUUAUAGAAGAGAGUGUG 60				
DB	101 GTTAAAGATGATCTTCTCTCTTAAACAATTTTGAAGAGTTATTAAGAGAAAGTACTG 160				
QY	61 CUAUCUUAUAUUGAGUUAUAUAUUGUUAUCUGUUCAGAGUCCUAUUGCAGGCC 120				
DB	161 CTATTTTAAATATTAAGTTAACTATTAATTTTACTGTTCAGATGCTTATTTGACAGCCC 220				
QY	121 CAUAUAUUCAGAGACACCCUCCUCCUUAUAUUGUUAUUGAGUGUCUAUUUAGAUAAGA 180				
DB	221 CATATATATCCAGAGACACCTCTCTGCTTTATATGATTAAGTTGCTATTTAAGATTAAGA 280				
QY	181 AAUAUAACCU 189				
DB	281 AAATTAACCT 289				
RESULT 4					
LOCUS	CO337794	539 bp	mRNA	linear	EST 05-OCT-2004
DEFINITION	EN15317.5prime Exelixis FlyTag MN08 Bluescript Drosophila melanogaster cDNA clone EN15317 5, mRNA sequence.				
ACCESSION	CO337794				
VERSION	CO337794.1	GI:49398069			
KEYWORDS	EST.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Archopoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
AUTHORS	1 (bases 1 to 539)				
TITLE					
JOURNAL					
COMMENT					
FEATURES					
ORIGIN					

DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
W601026.B21.1.D16 SS-R-N-A-11 Picea sitchensis cDNA clone W601026.D16 37, mRNA sequence.	C0224988	C0224988.1	GI:49047303 EST.	Picea sitchensis (Sitka spruce)	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.	Ralph,S., Kolosova,N., Oddy,C., Cooper,D., Butterfield,Y., Kirkpatrick,R., Liu,J., Palmquist,D., Scott,J., Barber,S., Yang,G., Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Maasson,A., Mayo,W., Moran,J., Olson,T., Wong,D., Friedmann,M.F., Ritland,C.E., Siddiqui,A., Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and Bohlmann,J. The spruce transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries Unpublished (2005) Contact: Joerg Bohlmann Genome BC forest genomics program University of British Columbia Michael Smith Laboratories, 6174 University Boulevard, Rm. 237, Vancouver, British Columbia, Canada, V6T 1Z3 Tel: 1-604-822-0282 Fax: 1-604-822-2114 Email: bohlmann@mel.ubc.ca plate: WS01026 row: D column: 16 High quality sequence stop: 448 POLVA=Yes.			

[illegible]

RESULT 8	
LOCUS	C2956789
DEFINITION	C2956789 873 bp DNA linear GSS 11-AUG-2005
ACCESSION	274903 Tomato EcorI BAC library Lycopersicon esculentum genomic
VERSION	clone_SL_EcorI0065f03_3, genomic survey sequence.
KEYWORDS	C2956789
SOURCE	C2956789.1 GI:72305252
ORGANISM	GSS.
REFERENCE	Lycopersicon esculentum (Solanum lycopersicum)
AUTHORS	Lycopersicon esculentum
TITLE	Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
JOURNAL	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
COMMENT	asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
	1 (bases 1 to 873)
	Mueller, J.A., Buehl, R.M., Wang, Y., Tanksley, S.D., Giovannoni, J.J.,
	Van Eck, J. and Stack, S.
	BAC end sequencing from three Solanum lycopersicon libraries
	Unpublished (2005)
	Contact: Lukas Mueller
	Tanksley Lab, Dept. of Plant Breeding
	Cornell University
	251 Emerson Hall, Ithaca, NY 14853, USA
	Tel: 607-255-6557
	Fax: 607-255-6683
	Email: sgn-feedback@sgn.cornell.edu
	Plate: 65 row: P column: 3
	Seq primer: SP6
	Class: BAC ends
	High quality sequence start: 99
	High quality sequence stop: 358.
FEATURES	
source	1..873 Location/Qualifiers

ORIGIN

Query Match	22.3%	Score 42.2	DB 10	Length 873
Best Local Similarity	29.9%	Pred. No. 0.51		
Matches	50	Conservative 39	Mismatches 78	Indels 0
			Gaps	0
QY	23	UUAUACAUAUUUGAGAGGUAUAUAGAAGAGAGUGAUCUUAUAUAUAGUUAAC	82	
Db	665	TTAATCCACTTGAATTAAACAACAATTGACTTATTTCATATATATTTGAAACTATA	724	
QY	83	UAUUUAGUUUAUCUGUUCAGAGUCCUAUUGSCAGCCCAUAUAUACGACACCCUUC	142	
Db	725	TATGAATTTGCTTTTCTTTTTTTTTCCTTAAGAAATTTTTTTTTTATATCACAATACCATT	784	
QY	143	CUGGCUUAUAUGAUAUAGUGUGCAUUAUAGAAUAUAACCU	189	
Db	786	TTGATTTTATATACGACGAGCTTTTCGTTAAATATATACATAATTCCT	831	

RESULT 9

LOCUS AZ748753 748 bp DNA 11mar GSS 25-JAN-2001
DEFINITION RPECI-24-65N21, TJ RPECI-24 Mus musculus genomic clone RPECI-24-65N21,
genomic survey sequence.
ACCESSION AZ748753
VERSION AZ748753.1 GI:12533912
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 748)
AUTHORS Zhao,S., Nierman,W., Malek,J., Shateman,S., Aktiret,B., Levins,M.,
Teegeye,G., Geer,K., Krol,W., Shartsbeyn,A., Gebregorjis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPCI-24
JOURNAL Unpublished (1999)
COMMENT Other GSSs: RPCI-24-65N21.TV
Contact: Shayng Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhaoc@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (http://www.choi.org/bacpac/orderingframe.html). BAC end
page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Seq primer: SP6
Class: BAC ends.
FEATURES
source location/Qualifiers
1..748
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-65N21"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/note="Vector: pTARBAC1, Site 1: BamHI, Site 2: BamHI,
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

ORIGIN
Query Match 21.2%; Score 40; DB 9; Length 748;
Best Local Similarity 29.9%; Pred. No. 2;
Matches 55; Conservative 39; Mismatches 90; Indels 0; Gaps 0;
QY 2 UUAAGUGUAGUUCUGUCUUAUACAUUUUGAGAGUUAUAAGAAGAAUAGUGC 61
DB 289 TTAAGATTGTGACAGAAATCTGAAAGCATTTTAATGAAATTGAGGTGTACATTTC 230
QY 62 UAUCUUAUAUUVAGUUAUCUUAUUGUACUUGCAGAGUCCUAGGACGCC 121
DB 229 AACGTAAACATCTACTTGTCTTTTGTGTAATGCCCTTATTATCTTTCGCAACAC 170
QY 122 AUAUAUUCGAGACAGACCCUUCUGUCUUAUUGAUUGUUGUCAUUAUAAGA 181
DB 169 AATGCACTATTACTAAATAATATGTGCTGTGTCATTAAATTCATCCAACTTAAA 110
QY 182 AUA 185
DB 109 AATA 106

RESULT 10
LOCUS CZ338852 942 bp DNA linear GSS 24-MAR-2005
DEFINITION ZMMBF0107A19F Zee may's genomic clone ZMMBF0107A19 5', genomic
survey sequence.
ACCESSION CZ338852
VERSION CZ338852
KEYWORDS GSS.
SOURCE Zee may's
ORGANISM Zee may's
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zee.

REFERENCE 1 (bases 1 to 942)
AUTHORS Bharti,A.K., Nelson,A.B., Young,S., Keizer,G., Zohovetz,V., Fuks,G.
and Messing,J.
TITLE Construction, Sequencing and Characterization of a Fosmid Library
JOURNAL of the B73 Maize Genome
Unpublished (2005)
CONTACT: Bharti,A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: T7
Class: fosmid ends.
FEATURES
source location/Qualifiers
1..942
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBF0107A19"
/lab_host="Epil100-T1"
/clone_lib="ZMMBF"
/note="Vector: pEpFOS-5; Site_1: Eco72I"

ORIGIN
Query Match 21.2%; Score 40; DB 10; Length 942;
Best Local Similarity 35.5%; Pred. No. 2.1;
Matches 54; Conservative 28; Mismatches 70; Indels 0; Gaps 0;
QY 1 GUUAAGUGUAGUUCUGUCUUAUACAUUUUGAGAGUUAUAAGAAGAAUAGUG 60
DB 402 GTTATATGCAAAAGATATGACATTAACAGTTGGAGATTAAATATATAATAGAT 461
QY 61 CUUUCUUAUAUUVAGUUAUCUUAUUGUACUUGCAGAGUCCUAGGACGCC 120
DB 462 CTACTCTATTAATTAATTTTAAAGCTTAGTTCATCTGCAAGTTACATGCGCTGGC 521
QY 121 CAUAUAUUCGAGACAGACCCUUCUGUCUUA 152
DB 522 GCGAAAACGAGGAGCAGCGGTTGATTCCTA 553

RESULT 11
LOCUS AG520404/c 1642 bp DNA linear GSS 22-DEC-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMG01-427C02.T7, genomic survey
sequence.
ACCESSION AG520404
VERSION AG520404.1 GI:48227817
KEYWORDS GSS.
SOURCE Mus musculus molossinus (Japanese wild mouse)
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
Abe,K., Noguchi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T.,
Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriwaki,K. and
Shiroishi,T.
Contribution of Asian mouse subspecies Mus musculus molossinus to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)
15574823
2 (bases 1 to 1642)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-Ku, Yokohama, Kanagawa, 230-0045, Japan


```

: INFORMATION FOR SEQ ID NO: 12
:
: SEQUENCE CHARACTERISTICS:
:
:   LENGTH: 4328
:   TYPE: nucleic acid
:   STRANDEDNESS: double
:   TOPOLOGY: linear
:
: US-08-322-742-12

```

Query Match	18.3%	Score 34.6;	DB 2;	Length 4328;
Best Local Similarity	27.2%;	Pred. No. 0.97;		
Matches 46;	Conservative 39;	Mismatches 84;	Indels 0;	Gaps 0;

[illegible]

RESULT 6
US-08-322-742-15/c
7. Sequence 15, Application US/08322742

GENERAL INFORMATION:
APPLICANT: Seger, Ruth
TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM ps/2 Model 502 or 555X
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,742

1 PRIOR APPLICATION DATA:
2 APPLICATION NUMBER: 07/938 8233
3 FILING DATE: September 1, 1992
4 APPLICATION NUMBER: 07/844,256
5 FILING DATE: February 28, 1992
6 APPLICATION NUMBER: 07/552,216
7 FILING DATE: February 28, 1991
8 ATTORNEY/AGENT INFORMATION:

1 REGISTRATION NUMBER: 34,819
2 REFERENCE/DOCKET NUMBER: 00530/048003?
3 TELECOMMUNICATION INFORMATION:
4 TELEPHONE: (617) 542-5070
5 TELEFAX: (617) 542-8906
6 TELEX: 200154
7 INFORMATION FOR SEQ ID NO: 15:
8 SEQUENCE CHARACTERISTICS:

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-322-742-15

Query Match	18.3%;	Score 34.6;	DB 2;	Length 4486;
Best Local Similarity	27.2%;	Pred. No. 0.98;		

Matches 46; Conservative 39; Mismatches 84; Indels 0; Gaps 0;

[illegible]

RESULT 7
US-09-949-016-15974
; Sequence 15974, Application US/09949016

```

1  APPLICANT: VENTER, J. Craig et al.
2  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
3  TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
4  FILE REFERENCE: CL001307
5  CURRENT APPLICATION NUMBER: US/09/949,016
6  CURRENT FILING DATE: 2000-04-14
7  PRIOR APPLICATION NUMBER: 60/241,755
8  PRIOR FILING DATE: 2000-10-20
9  PRIOR APPLICATION NUMBER: 60/237,768
10 PRIOR FILING DATE: 2000-10-03
11 PRIOR APPLICATION NUMBER: 60/231,498
12 PRIOR FILING DATE: 2000-09-08
13 NUMBER OF SEQ ID NOS: 207012
14 SOFTWARE: FastSeq for Windows Version 4.0
15 SEQ ID NO 15974
16 LENGTH: 294836
17 TYPE: DNA
18 ORGANISM: Human
19 FEATURE:
20 NAME/KEY: misc_feature
21 LOCATION: (1)..(294836)
22 OTHER INFORMATION: n = A,T,C or G
23 JS-09-949-016-15974

```

Query Match	17.2%;	Score 32.6;	DB 3;	Length 294836;
Best Local Similarity	31.0%;	Pred. No. 14;		
Matches 27;	Conservative 26;	Mismatches 34;	Indels 0;	Gaps 0;

[illegible]

```

RESULT 8
US-09-949-016-17543
; Sequence 17543; Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```



```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 17543
; LENGTH: 36546
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17543

Query Match      17.0%, Score 32.2, DB 3, Length 36546;
Best Local Similarity 32.7%, Pred. No. 10;
Matches 33, Conservative 25, Mismatches 43, Indels 0, Gaps 0;

Qy      15  UGCUUCUCCUUAACUAAUUUGAGAGUUAUAAGAAGAGAAUAGUCUACUUAUAU 74
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      22615 TTACTTCTACTTAAAGATCACTTGAGTTCATGAGAGAAGAGACAGTCATTTGTCTGT 22674

Qy      75  AGGUUAACUUAUUAGCUUUUACUCUUCAGAGUCCCUAUTC 115
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      22675 AACCTAATTCACCAATTTTAAATATGCGCAATGATAGGCG 22715

```

```

RESULT 9
US-08-916-421B-2
; Sequence 2, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschi
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 58407
; TYPE: DNA
; ORGANISM: Methanococcus jannaschi
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (6485)..(6485)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-2

```

[illegible]

RESULT 10
US-09-692-570-2
; Sequence 2, Application US/09692570
; Patent No. 6797466

```

: GENERAL INFORMATION:
: APPLICANT: Bult et al.
: TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
: Patent No. 6797466
: TITLE OF INVENTION: jannaschl1
: FILE REFERENCE: P3275C1
: CURRENT APPLICATION NUMBER: US/09/692,570
: CURRENT FILING DATE: 2003-01-14
: PRIOR APPLICATION NUMBER: US 60/024,428
: PRIOR FILING DATE: 1996-08-22
: PRIOR APPLICATION NUMBER: US 08/916,421
: PRIOR FILING DATE: 1997-08-22
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: Patentin version 3.1
: SEQ ID NO 2
: LENGTH: 58407
: TYPE: DNA
: ORGANISM: Methanococcus jannaschl1
: FEATURES:
: NAME/KEY: misc feature
: LOCATION: (6485)..(6485)
: OTHER INFORMATION: n equals a, t, c, or g
: US-09-692-570-2

```

[illegible]

```

RESULT 11
US-09-949-016-17398/C
; Sequence 17398, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 17398
; LENGTH: 139150
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-17398

```

Query Match	17.0%;	Score 32.2;	DB 3;	Length 139150;
Best Local Similarity	23.7%;	Pred. No. 15;		
Matches	22;	Conservative	33;	Mismatches 38;
				Indels 0;
				Gaps 0;


```

Qy      7  UGUGAUCUUCUUCUUAACAUAUUUGAGAGUUAUAAGAAGAAUGAGCUAUCU 66
Db      17606  AAGTTTTTATTTCTCTTAAATAATTTGACCTGTATTAAAAAAGCAGGTATTATT 17547
Qy      67  UAAUUAUVAGGUUAACUUAUUUGUUVUUAUCUGU 99
Db      17546  TATTTATTTTATTTGTTTTTAAATTTTGAAGATT 17514

RESULT 12
US-09-949-016-12879/c
; Sequence 12879, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12879
; LENGTH: 139577
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12879

Query Match      17.0%; Score 32.2; DB 3; Length 139577;
Best Local Similarity 23.7%; Pred. No. 15;
Matches 22; Conservative 33; Mismatches 38; Indels 0; Gaps 0;

Qy      7  AUGGCAUCUUCUUCUUAUAACAUAUUUGAGAGUUAUAAGAAGAAUGAGCUAUCU 66
Db      17606  AAGTTTTTATTTCTCTTAAATAATTTGACCTGTATTAAAAAAGCAGGTATTATT 17547
Qy      67  UAAUUAUVAGGUUAACUUAUUUGUUVUUAUCUGU 99
Db      17546  TATTTATTTTATTTGTTTTTAAATTTTGAAGATT 17514

RESULT 13
US-09-949-016-12928
; Sequence 12928, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12928
; LENGTH: 194537
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: (1)...(194537)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12928

Query Match      17.0%; Score 32.2; DB 3; Length 194537;
Best Local Similarity 34.2%; Pred. No. 17;
Matches 40; Conservative 24; Mismatches 53; Indels 0; Gaps 0;

Qy      17  GCUUCCUUAUACAUVUUUGAGAGGCUUAUUAAGAGAGAAUGAGUCUACUUAUAUUAUUG 76
      ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      163390 GCTTACACATGTACCAATTAAGAAAGTTTAATTAATAAACAGATTAATGTTTACTCAATGTGG 176449

Qy      77  GUUAACUUAUUAGUUUUUACUGUUCAGAGUCCUUAUUGCAGGCCCAUAUAUCCAG 133
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      164450 GTGAATTAAGTAGAGTGATAGCAGTCATGGTGTGGGTTTAATCAAAAGTATTAATG 176506

RESULT 14
US-09-949-016-12740
; Sequence 12740, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12740
; LENGTH: 201529
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(201529)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12740

Query Match      17.0%; Score 32.2; DB 3; Length 201529;
Best Local Similarity 34.2%; Pred. No. 17;
Matches 40; Conservative 24; Mismatches 53; Indels 0; Gaps 0;

Qy      17  GCUUCCUUAUACAUVUUUGAGAGGCUUAUUAAGAGAGAAUGAGUCUACUUAUAUUAUUG 76
      ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      163390 GCTTACACATGTACCAATTAAGAAAGTTTAATTAATAAACAGATTAATGTTTACTCAATGTGG 163449

Qy      77  GUUAACUUAUUAGUUUUUACUGUUCAGAGUCCUUAUUGCAGGCCCAUAUAUCCAG 133
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      164450 GTGAATTAAGTAGAGTGATAGCAGTCATGGTGTGGGTTTAATCAAAAGTATTAATG 163506

RESULT 15
US-09-299-378-3
; Sequence 3, Application US/09299378
; Patent No. 6245545
; GENERAL INFORMATION:
; APPLICANT: Higgins, Lauren S.
; APPLICANT: Dalton, Michael A.
; TITLE OF INVENTION: Method For Cloning And Producing The Swai Restriction
; FILE REFERENCE: Swai
; CURRENT APPLICATION NUMBER: US/09/299,378
; CURRENT FILING DATE: 1999-04-27
; NUMBER OF SEQ ID NOS: 20

```



```

US-10-088-750B-4
; Sequence 4, Application US/10088750B
; Publication No. US20040166486A1
; GENERAL INFORMATION:
; APPLICANT: NAKASHIMA, Nobuhiko
; APPLICANT: KANAMORI, YASUSHI
; TITLE OF INVENTION: A Novel Higher-Order Structure With Promoting Translation
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 3190-015
; CURRENT APPLICATION NUMBER: US/10/088, 750B
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: JP P2001-016746
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/JP01/00641
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 200
; TYPE: RNA
; ORGANISM: Cricket Paralysis Virus
; US-10-088-750B-4

```

	Query Match	65.4%	Score 123.6	DB 7	Length 200
	Best Local Similarity	79.0%	Pred. No. 2.4e-25		
	Matches 147	Conservative 0	Mismatches 39	Indels 0	Gaps 0
QY	4	AAAGATGAGACGCTGCTCCCTGAAACAATTTGAGAGAGTAAATGAAGAAGAAAGTATGCTCA	63		
Db	3	AAAATGUGAGCTGCTGCTGGAATTAATATTTGAGAGCTGAAATTAATTAACAATGATGCTCA	62		
QY	64	UCCTTAATTAATGAGTAAACGATTTGATGCTTAACTGTTGAGATGCTCAATGACAGCCCTCA	123		
Db	63	TTTTTGATTTATGAGTAACTGATATTTGATGCTTAACTGTTGAGATGCTCAATGACAGCCCTCA	122		
QY	124	AAATATCCAGGACACACCTCTCTGCTCTCTATTAATGATTTGAGTGTGATTTGAAATGAAAA	183		
Db	123	AAATATCCAGGACAGCCCTCTCTGCTGCTGCTTATGATTTGAGTGTGATTTGAAAAATCCATGAAAA	182		
QY	184	TAACCT	189		
Db	183	TTTCTT	188		

```

// RESULT 3
// US-10-088-750B-6
// Sequence 6, Application US/10088750B
// Publication No. US20040166486A1
// GENERAL INFORMATION:
// APPLICANT: NAKASHIMA, Nobuhiko
// APPLICANT: KANAOKA, Yasushi
// TITLE OF INVENTION: A Novel Higher-Order Structure With Promoting Translation
// FILE REFERENCE: 3190-015
// CURRENT APPLICATION NUMBER: US/10/088,750B
// CURRENT FILING DATE: 2002-03-20
// PRIOR APPLICATION NUMBER: JP P2001-016746
// PRIOR FILING DATE: 2001-01-25
// PRIOR APPLICATION NUMBER: PCT/JP01/00641
// PRIOR FILING DATE: 2001-01-31
// NUMBER OF SEQ ID NOS: 12
// SOFTWARE: PatentIn version 3.2
// SEQ ID NO 6
// LENGTH: 202
// TYPE: RNA
// ORGANISM: Black Queen-Cell Virus
// US-10-088-750B-6

```

7 AUGUGAUCUUGCUUCCUUAUAC-AAUUUGAGAGGUAUAAGAGGAGUAGUGCUAUC 65

Db 7 AUGUGAUCUUGUCUGUGGAGGCGAAAUUUUGCA CAAUUAUAAAUUUCGACAAUGAUGGCUAAU 66

Qy 66 UUAUUAUUAGGUAUACUAUUUAAGUUUUAUCUGUUUCAGAGUGCCUAU -UGGCAAGCCCCAUA 124

Db 67 GUUGGAAUUCACCGUACUAUUUAAGGUUAACGUCACAAAUUCGUGAGAAAGCAGCCCUAUC 126

Qy 125 AUUUAUCCAGGACACCCUUCUGUCUUAUUAUGAUUAGGUGUUCAUUUUUGAUUAUAGAAU 184

Db 127 AAUUUUAAGAGAAUCUGUGCUAUGUUUUUGAAGAAUUAUGUAUUCUUAACAGAAACAUUU 186

Qy 185 AAC 187

Db 187 ACC 189

```

; RESULT 4
; US-11-097-143-25048/c
; Sequence 25048, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: C1000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 25048
; LENGTH: 5059
; TYPE: DNA
; ORGANISM: DROSOPHILA
; US-11-097-143-25048

```

Query Match	21.0%	Score 39.6	DB 10	Length 5059
Best Local Similarity	31.3%	Pred. No. 1		
Matches	57	Conservative 36	Mismatches 89	Indels 0
			Gaps	0
QY	2	UUAAGUGUGAUCUCUCUCUUAACAUNUGAGAGUUAUAAAGAAAGAAUCUC	61	
	::: :	::: :		
Db	661	TTAAATATAAAAACAATTAGTACATATTTTAAAGAAAAAGAAATATATTTCTTAA	602	
QY	62	UUUCUUAUUUAUUGAGUUAACUUAUUUGUUUUAUCUGUCAGAUCCUUAUUGCAGACCCC	121	
	::: :	::: :		
Db	601	TATCGAAGTACTATATATATATCTTTCTTACCTAAAAAAAACAATAGATTTCTCTA	542	
QY	122	AUAUAUUCGACGACACCCUUCUGUCUCUUAUAUGAUAUGUGUGUCUUAUAUAAGA	181	
	::: :	::: :		
Db	541	TAAACATCAATGTTAACTTGTTGCTGCTAACTTTTAATGTTATCTATCGAAAAACAA	482	
QY	182	AA 183		
Db	481	AA 480		

Db 128944 TTAGTCTTAGAGCTTATGAGCTTATGAATTTGTAAAGAGACTTTTTCATTT 129003
QY 152 AUUU 155
Db 129004 ATTT 129007

RESULT 12
US-10-195-144-87
; Sequence 87, Application US/10195144
; Publication No. US20030126646A1
; GENERAL INFORMATION:
; APPLICANT: BROWN, GREGORY G.
; APPLICANT: FORMANOVA, NATASA
; APPLICANT: DENDY, CHARLES
; APPLICANT: LANDRY, BENOIT S.
; APPLICANT: CHEUNG, WING
; APPLICANT: JIN, HUA
; TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 16313-0136
; CURRENT APPLICATION NUMBER: US/10/195.144
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/305,026
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/305,363
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/308,736
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 271990
; TYPE: DNA
; ORGANISM: Raphanus sativum
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (144241)..(144300)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-195-144-87

Query Match 18.5%; Score 35; DB 6; Length 271990;
Best Local Similarity 37.3%; Pred. No. 97;
Matches 28; Conservative 22; Mismatches 25; Indels 0; Gaps 0;

QY 19 UUCUUUACAUUUGAGAGUUAAUAGAAGAGUAGUCUUAUUAUUAGU 78
Db 97686 TTGCATATTATTTATTAATCTTCAAAAGATGAAGTAGCTGTAGATAATTAAAG 97745
QY 79 UACUUAUUAGUUU 93
Db 97746 TGACAACTTAAAT 97760

RESULT 13
US-10-345-072-87
; Sequence 87, Application US/10345072
; Publication No. US20030237112A1
; GENERAL INFORMATION:
; APPLICANT: BROWN, GREGORY G.
; APPLICANT: FORMANOVA, NATASA
; APPLICANT: DENDY, CHARLES
; APPLICANT: LANDRY, BENOIT S.
; APPLICANT: CHEUNG, WING
; APPLICANT: JIN, HUA
; APPLICANT: LAI, PANG MING
; APPLICANT: LEROBERT, MARTIN
; TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 16313-0210
; CURRENT APPLICATION NUMBER: US/10/345.072
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: PCT/US02/22217

; PRIOR FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 60/305,026
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/305,363
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/308,736
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 271990
; TYPE: DNA
; ORGANISM: Raphanus sativum
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (144241)..(144300)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-345-072-87

Query Match 18.5%; Score 35; DB 6; Length 271990;
Best Local Similarity 37.3%; Pred. No. 97;
Matches 28; Conservative 22; Mismatches 25; Indels 0; Gaps 0;

QY 19 UUCUUUACAUUUGAGAGUUAAUAGAAGAGUAGUCUUAUUAUUAGU 78
Db 97686 TTGCATATTATTTATTAATCTTCAAAAGATGAAGTAGCTGTAGATAATTAAAG 97745
QY 79 UACUUAUUAGUUU 93
Db 97746 TGACAACTTAAAT 97760

RESULT 14
US-10-433-793-108
; Sequence 108, Application US/10433793
; Publication No. US20040142334A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von mit Angiogenese assoziierten Krankheiten
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/433.793
; CURRENT FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 212
; SEQ ID NO 108
; LENGTH: 37184
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-433-793-108

Query Match 18.4%; Score 34.8; DB 7; Length 37184;
Best Local Similarity 32.4%; Pred. No. 52;
Matches 33; Conservative 27; Mismatches 42; Indels 0; Gaps 0;

QY 16 UGCUUCCUUAACAUAUUUGAGAGUUAAUAGAAGAGUAGUCUUAUUAUU 75
Db 27976 TGGTTATTAGTTTATTTTGGATGATGATGAATTTAGGATTGGTTTGAATATTA 28035
QY 76 GGUUACUUAUUAGUUUACUGUUCAGAGUCCUUAUUGCAG 117
Db 28036 GGTTAAGTATTAGGTATACGGGTAAATGAAATTTTGGAGG 28077

RESULT 15
US-10-312-841-1
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/MO
; CURRENT APPLICATION NUMBER: US/10/312.841

; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match 18.1%; Score 34.2; DB 6; Length 3673778;
Best Local Similarity 20.7%; Pred. No. 4e+02;
Matches 23; Conservative 40; Mismatches 48; Indels 0; Gaps 0;

QY 2 UUAAGUGUGAUCUCCUUAUACAUAUUUGAGAGUUAUAGAGAGAGUAGUC 61
DB 2429357 TAAGGATTTTTCGTATTTATTATTTAATTTAATTTAATTTAATTTAATTTA 2429416
QY 62 UAUCUUAUAUAUAGUUAACUAVUUAAGUUUACUGUUCAGAGUCCUAVU 112
DB 2429417 TATGATTAAGTATTAGTTATTTTATTTTATTTTATTTTATTTTATTTAT 2429467

Search completed: December 22, 2005, 16:21:06
Job time : 446.936 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	35.2	18.6	994	6	US-10-750-185-25532	Sequence 25532, A
2	32.8	17.4	941	6	US-10-750-185-54894	Sequence 54894, A
3	32.4	17.1	176503	7	US-11-121-086-53	Sequence 53, Appl
4	32.4	17.1	1082144	7	US-11-117-187-211	Sequence 211, Appl
5	32.2	17.0	201	6	US-10-995-561-65121	Sequence 65121, A
6	32.2	17.0	1358	6	US-10-750-185-51890	Sequence 51890, A
7	32.2	17.0	66916	6	US-10-750-185-13374	Sequence 13374, A
8	32	16.9	1732	6	US-10-750-185-34293	Sequence 34293, A
9	31.8	16.8	3000	6	US-10-750-185-26975	Sequence 26975, A
10	31.8	16.8	3824	6	US-10-750-185-47663	Sequence 47663, A
11	31.4	16.6	1131	6	US-10-750-185-47659	Sequence 47659, A
12	31	16.3	1210	6	US-10-750-185-29599	Sequence 29599, A
13	30.8	16.3	96583	7	US-11-117-187-203	Sequence 203, Appl
14	30.6	16.2	321	6	US-10-793-626-1161	Sequence 1161, Appl
15	30.6	16.2	2170	6	US-10-750-185-33762	Sequence 33762, A
16	30.6	16.2	2236	6	US-10-750-185-39368	Sequence 39368, A
17	30.6	16.2	3076	6	US-10-793-626-3643	Sequence 3643, Appl
18	30.6	16.2	3228	6	US-10-793-626-3808	Sequence 3808, Appl
19	30	15.9	993	6	US-10-750-185-60349	Sequence 60349, A
20	30	15.9	1417	6	US-10-750-185-48773	Sequence 48773, A
21	30	15.9	2031	6	US-10-750-185-39840	Sequence 39840, A
22	30	15.9	2251	6	US-10-750-185-5678	Sequence 64707, A
23	30	15.9	150468	7	US-11-112-908-56	Sequence 56, Appl

C 24	30	15.9	193789	7	US-11-112-908-55	Sequence 55, Appl A
C 25	29.8	15.8	2051	6	US-10-750-185-4521	Sequence 4531, A
C 26	29.6	15.7	7785	6	US-10-955-05A-63	Sequence 65, Appl A
C 27	29.9	15.7	159188	6	US-10-995-561-1330	Sequence 13230, A
C 28	29.6	15.6	309	6	US-10-793-626-357	Sequence 357, Appl A
C 29	29.4	15.6	309	6	US-10-793-626-1407	Sequence 1407, Appl A
C 30	29.4	15.6	1193	6	US-10-793-626-1421	Sequence 1421, Appl A
C 31	29.4	15.6	1239	6	US-10-793-626-367	Sequence 367, Appl A
C 32	29.4	15.6	1357	6	US-10-750-185-4563	Sequence 45753, A
C 33	29.4	15.6	1894	6	US-10-793-626-4364	Sequence 4366, A
C 34	29.4	15.6	2132	6	US-10-485-511-398	Sequence 398, Appl A
C 35	29.4	15.6	2760	6	US-10-750-185-52376	Sequence 52376, A
C 36	29.4	15.6	3021	6	US-10-793-626-3318	Sequence 4315, Appl A
C 37	29.4	15.6	3285	6	US-10-793-626-3728	Sequence 3728, A
C 38	29.4	15.6	3758	6	US-10-793-626-4145	Sequence 4145, Appl A
C 39	29.2	15.4	1222	6	US-10-750-185-55775	Sequence 55775, A
C 40	29.2	15.4	1512	6	US-10-750-185-42903	Sequence 42903, A
C 41	29.2	15.4	1655	6	US-10-750-185-31674	Sequence 31674, A
C 42	29.2	15.4	2968	6	US-10-750-185-47496	Sequence 47496, A
C 43	29	15.3	201	6	US-10-995-561-18111	Sequence 18111, A
C 44	29	15.3	201	6	US-10-995-561-22013	Sequence 22013, A
C 45	29	15.3	1256	6	US-10-750-185-35396	Sequence 35396, A

ALIGNMENTS

```

RESULT 1
US-10-750-185-25522
; Sequence 25522, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FATTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25522
; LENGTH: 994
; TYPE: DNA
; ORGANISM: Bovine 19866881820768
US-10-750-185-25522

Query Match 18.6%; Score 35.2; DB 6; Length 994;
Best Local Similarity 27.7%; Pred. No. 0.42;
Matches 31; Conservative 33; Mismatches 48; Indels 0; Gaps 0

QY 54 AGUAGUCUUVUUAUAUAGGUVUACUUAUAGUUVUACUGUCAGAUCCUUAUG 113
   |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 615 AGCGAGCATGCTTCATCATCTATGTGTGCTTAATTAATTAATCCAGTCTTACCTCG 674

QY 114 GCAGCCCCAUAUAUAVUCAGAGCACCCUCUCUCUCUUAUAUAGUUAAGUAG 165
   |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 675 TCAGCGTAGTGTGCGCTCTTACGTCTGTGTGTTTAATTAATTAATCCCTTGG 726

RESULT 2
US-10-750-185-54894/c
; Sequence 54894, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.

```



```

; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54894
; LENGTH: 941
; TYPE: DNA
; ORGANISM: Bovine 19866860677518
; US-10-750-185-54894
```

```
Query Match      17.4%; Score 32.8; DB 6; Length 941;
Best Local Similarity 27.9%; Pred. No. 2;
Matches 39; Conservative 34; Mismatches 67; Indels 0; Gaps 0;
```

```
QY      2 UDAAGUGUAGUCUCCUUAUACAUVUUGAGGUAUAAGAAGAGUAGUC 61
DB      371 TTAATAATGATTATTAATTAATGATGATGATTAAGATTAGTATCTGATGT 312

QY      62 UAUCAUAUAUUGAGUACUUAUUGUACUGUCAGAGGCUUAUUGGAGCC 121
DB      311 AAATGCCCTCAATATTAATGATGATGATTTACTTTTAACATTTTCACTGCTCA 252

QY      122 AUAUAUCCAGACACCCUC 141
DB      251 ATAAATTAAGCAACGCTC 232
```

```

RESULT 3
; US-11-121-086-53/c
; Sequence 53, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 53
; LENGTH: 176503
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-121-086-53
```

```
Query Match      17.1%; Score 32.4; DB 7; Length 176503;
Best Local Similarity 33.7%; Pred. No. 11;
Matches 33; Conservative 24; Mismatches 41; Indels 0; Gaps 0;
```

```
QY      1 GUAAGAUGUAGUCUCCUUAUACAUAUUGAGGUAUAAGAAGAGUAGUC 60
DB      99324 GCTATTCGCGATACGGTTTATACATATAGTTTGATGAAGAATGAAGATAAG 99265

QY      61 CUUAUUAUAUUGAGUACUUAUUGUUAUUGUUAUUGUUAUUGUUAUUG 98
DB      99264 AACTTTAAATTAAGATATTATATATCTCATAGT 99227
```

```

RESULT 4
; US-11-117-187-211
; Sequence 211, Application US/11117187
```

```

; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: FREUS, DAPHNE
; APPLICANT: COPELHAY, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 1082144
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-11-117-187-211
```

```
Query Match      17.1%; Score 32.4; DB 7; Length 1082144;
Best Local Similarity 30.9%; Pred. No. 17;
Matches 50; Conservative 31; Mismatches 81; Indels 0; Gaps 0;
```

```
QY      25 AUAUAUUGAGAGGUAUAUAAGAAGAGUAGUCUUAUAUAUUGAGUUAUAUA 84
DB      487526 AAACAATTAATGAGAGTTTCCAGGTTGTATCTGATGTTTGCAGAAATTAATCAAAATCTA 487585

QY      85 UUAUUAUUAUUGUACAGAGGCUUAUUGGACGCCCAUAUAUUGACAGACCCUCUCU 144
DB      487586 TTATCAATTTCTGATATATCAACGGTATCAAGATATCTTATTAACATTAACAAATTAAT 487645

QY      145 GCUUUAUAUUGAUUGUUGUACUUAUAUAAGAUAUAUAUAUA 186
DB      487646 TCCTGCTGCTTAATCAAGATATCTTACTTATTTATTA 487687
```

```

RESULT 5
; US-10-995-561-65122/c
; Sequence 65122, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CU001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65122
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-995-561-65122
```

```
Query Match      17.0%; Score 32.2; DB 6; Length 201;
Best Local Similarity 30.8%; Pred. No. 1.9;
Matches 36; Conservative 28; Mismatches 53; Indels 0; Gaps 0;
```

```
QY      10 UGAUUCUUGUCUUAUAUAUAUUGAGGUAUAUAAGAAGAGUAGUCUUAUAUA 69
DB      172 TATATTTCTAATATTCATCAACATTTGAAGATATATCTTGAATGATGAATTAAGAAA 113

QY      70 UUAUUGUUAUAUAUAUUAUUGUUAUUGGAGAGCCCAUAUAUUGGAGCCCAUAUAU 126
DB      112 ATATGATTTATCTGTTCTTATCTTATATATGCTACTAGAAAGCTTAATAT 56
```

```

RESULT 6
; US-10-750-185-51890/c
; Sequence 51890, Application US/10750185
```



```
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: PANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 51890
LENGTH: 1358
TYPE: DNA
ORGANISM: Bovine 19866881423804
US-10-750-185-51890

Query Match 17.0%; Score 32.2; DB 6; Length 1358;
Best Local Similarity 28.7%; Pred. No. 3.3;
Matches 45; Conservative 34; Mismatches 78; Indels 0; Gaps 0;

Qy 23 UUAUACAUAUUUGAGAGUUAUAGAAGUAGUCUUAUUAUUAUUGGUUAAC 82
Db 927 TTTTAGAGAGATTAAGTACTTAAATGTTGTGGGATTTATTAATACCAATGCAAT 868
Qy 83 UAUUAGUUUUAUCUGUAGAGUUGCCUUAUUGGACGCCCAUAUUAUUGCAGACCCU 142
Db 867 TTAATTTCTTTAAGCTTAAGATCTTTTGGCAAAATGCTCCTAGGAAACCATC 808
Qy 143 CUGCUUUAUUAUAGUAGUUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 179
Db 807 AATTCTCTTAATCATCTGAATCTCTTAATATACAG 771

RESULT 7
US-10-995-561-13374/c
Sequence 13374, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
FILE REFERENCE: CLO01559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13374
LENGTH: 66916
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1) ... (66916)
OTHER INFORMATION: n = A, T, C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13374

Query Match 17.0%; Score 32.2; DB 6; Length 66916;
Best Local Similarity 30.8%; Pred. No. 9.4;
Matches 36; Conservative 28; Mismatches 53; Indels 0; Gaps 0;

Qy 10 UGAUCUGUCUCCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 69
Db 51752 TAAATTTCTTAATCTTAATCAATGATTAATCTTAATCAATGATTAATCAATG 51693
Qy 70 UAUUAGUUAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 126
```

```
Db 51692 AATATGATTTATVCGTTTCTTACCTTTAATATGCGTACTGAAGCTCTAAAT 51636

RESULT 8
US-10-185-34293
Sequence 34293, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: PANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 34293
LENGTH: 1732
TYPE: DNA
ORGANISM: Bovine 19866880882569
US-10-750-185-34293

Query Match 16.9%; Score 32; DB 6; Length 1732;
Best Local Similarity 30.4%; Pred. No. 4;
Matches 56; Conservative 33; Mismatches 95; Indels 0; Gaps 0;

Qy 4 AAGAUGUAGUUCUUCUUAUAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 63
Db 906 AAGAAGGCCGAGCGCTCATCATATATTTGGAATTAATGAAGGGAAGGTTGTG 965
Qy 64 UCUAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 123
Db 966 TAAATTAAGTAATCTGAATATGTAATTAACAAGTACAAAGTATTAATTCACAGGACAT 1025
Qy 124 AAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 183
Db 1026 AAATATGACAGAAATCTCCCTTGATATGTTACTTACGAAATATTATGATATA 1085
Qy 184 UAAC 187
Db 1086 AGAC 1089

RESULT 9
US-10-750-185-26975
Sequence 26975, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: PANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 26975
LENGTH: 3000
```



```
; TYPE: DNA
; ORGANISM: Bovine 19866880583819
US-10-750-185-26975

Query Match      16.8%; Score 31.8; DB 6; Length 3000;
Best Local Similarity 32.7%; Pred. No. 5.2;
Matches 35; Conservative 25; Mismatches 47; Indels 0; Gaps 0;

Qy      83 UAUUUAGUUUUUACUUCUUGACGAGUCUUAUGGACCCCAUAUUAUCCAGACACCCUUC 142
Db      612 TGTTCAGTTCCTAACTCATGTCTGACTCTCTTGCAACCCCAAGGACTGCAGCACACGAGC 671

Qy      143 CUGCUUCUUAUUGAUUGUGUCUUAUUGAUUAGAUUAGAAUUAACCU 189
Db      612 GTCCTTGTGTGGGAAATCACTGTAATTCAATTTAAAAATCTTCT 718

RESULT 10
US-10-750-185-47663/c
; Sequence 47663, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47663
; LENGTH: 3824
; TYPE: DNA
; ORGANISM: Bovine 19866880747157
US-10-750-185-47663

Query Match      16.8%; Score 31.8; DB 6; Length 3824;
Best Local Similarity 36.1%; Pred. No. 5.6;
Matches 30; Conservative 21; Mismatches 32; Indels 0; Gaps 0;

Qy      34 UGAGAGUUAUUAAGAAGAGUAGUCUUAUCUUAUUAUAGUUAUCUUAUUUUAUUU 93
Db      1265 TTAGTAAAGTAAGATTAAACAAATCTGTCAATTTAAAGTAAGATTAACTTAAACTT 206

Qy      94 ACUGUUCAGAGUCCUUAUUGGCA 116
Db      1205 ACTCATTAAGCATGGTTTGTGA 1183

RESULT 11
US-10-750-185-47569/c
; Sequence 47569, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
```

```
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47569
; LENGTH: 1131
; TYPE: DNA
; ORGANISM: Bovine 19866881500383
US-10-750-185-47569

Query Match      16.6%; Score 31.4; DB 6; Length 1131;
Best Local Similarity 26.7%; Pred. No. 5.2;
Matches 28; Conservative 31; Mismatches 46; Indels 0; Gaps 0;

Qy      8 UUGCAUUCUUCGUCCUUAUUAUUAUUGAGAGUUAUUAAGAAAGUAGUGUCUUAUCU 67
Db      1102 TGTGTTTTTTTTTTTACAAATGCTGAGTTAAATGAATCTTTGACTGTGTCTTTT 1043

Qy      68 AAUAUUGAGUUAACUUAUUAUUGUUAUCUGUUCAGAGUCCUUAU 112
Db      1042 TATATTAGGTGAGAAATCATAGTTTAACTTTAAAAAGCCTCTT 998

RESULT 12
US-10-750-185-29599/c
; Sequence 29599, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29599
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Bovine 19866880922522
US-10-750-185-29599

Query Match      16.4%; Score 31; DB 6; Length 1210;
Best Local Similarity 26.3%; Pred. No. 6.9;
Matches 46; Conservative 39; Mismatches 90; Indels 0; Gaps 0;

Qy      9 GUGAUUCUUCGUCCUUAUUAUUAUUGAGAGUUAUUAAGAAAGUAGUGUCUUAUCU 68
Db      1011 GTGAAATGCTGAGATCATACAGTGTCTATGTTTAACTTTGAAGACCTGCAACTA 952

Qy      69 AUAUUGAGUUAACUUAUUAUUGUUCAGAGUCCUUAUUGGACCCCAUAUUAU 128
Db      951 TTTTCAACAAGCTGTACATTATTAATTCCTGAGCAAGCAAGGTTGAATT 892

Qy      129 CCAGACACCCUUCUUCGUCCUUAUUAUUGAGUUGUUCUUAUUAAGAAAA 183
Db      891 TCTCCATATCCCTCCAAATACATTAATTTCTGGGTTTGTCTTGATATA 837

RESULT 13
US-11-117-187-203
; Sequence 203, Application US/1117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: PREUS, DAPHNE
; APPLICANT: COPEHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
```



```
FILE REFERENCE: ARCD:109US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 203
; LENGTH: 96583
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-11-117-187-203

Query Match      16.3%; Score 30.8; DB 7; Length 96583;
Best Local Similarity 34.9%; Pred. No. 26;
Matches 37; Conservative 22; Mismatches 47; Indels 0; Gaps 0;

Oy      81 ACUAAUUAGUUUACUGUUCAGAGCCGUAUUGGACGCCCAUAUAUACGACGACCCU 140
Db      70104 ACTTTCTCTTGAACTTTAGCAAAACCGATTCTCAGATCATCAAACTACTGATCT 70163
Oy      141 CUCUGCUCUUAUUAUGAUUAGUGUCUAUUGAUAUAAGAAUAUA 186
Db      70164 CCGTCGACCCATCCGATTGGATTGATTATCTGTAACAAAGTAA 70209

RESULT 14
US-10-793-626-1161/c
; Sequence 1161, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1161
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1161

Query Match      16.2%; Score 30.6; DB 6; Length 321;
Best Local Similarity 32.3%; Pred. No. 6.2;
Matches 30; Conservative 24; Mismatches 39; Indels 0; Gaps 0;

Oy      2 UUAAGAGUGAGUUCUGUCUCCUUAUACAUAUUUGAGAGUUAUAAGAAGAGUAGUC 61
Db      120 TTCAGCTATGGTCATATATCTGGAGATATTTTGTACAAAGTAATACATATCTATCTTC 61
Oy      62 UAUUUUAUAUAUUAUGUUAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 94
Db      60 TTTTATAATATATATGACAGAACATTAATTCTAA 28

RESULT 15
US-10-750-185-33762
; Sequence 33762, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
```

```
APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 33762
; LENGTH: 2170
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-33762

Query Match      16.2%; Score 30.6; DB 6; Length 2170;
Best Local Similarity 26.6%; Pred. No. 10;
Matches 29; Conservative 31; Mismatches 49; Indels 0; Gaps 0;

Oy      6 GAUGUGAUCUGUCUCCUUAUACAUAUUUGAGAGUUAUAAGAAGAGUAGUCUUAUC 65
Db      262 GAATTTATTTATTTATGATCTCTGAGATCTTTTAAGACATACAGTCTTATA 321
Oy      66 UUAUAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 114
Db      322 TTCAATTTAGATATCTAATTTATTCAGTTTATGAGATGATTTTAA 370
```

Search completed: December 22, 2005, 16:41:03
Job time : 175.462 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 04:46:32 ; Search time 1483.08 Seconds
(without alignments)
7205.664 Million cell updates/sec

Title: US-10-088-750C-4

Perfect score: 188
Sequence: 1 caaaagaugauucugcug.....aaaacuaagaauuacuu 188

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl1:
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_ov:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_dr:*
9: gb_to:*
10: gb_str:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hcg:*
15: gb_dl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	188	100.0	188 6 BD173514	BD173514 Novel ter
2	188	100.0	188 6 BD177018	BD177018 Novel ter
3	188	100.0	9185 13 AF218039	AF218039 Cricket p
4	125.2	66.6	1345 6 AR496112	AR496112 Sequence
5	125.2	66.6	1345 6 AR511394	AR511394 Sequence
6	123.6	65.7	189 6 BD173513	BD173513 Novel ter
7	123.6	65.7	189 6 BD177017	BD177017 Novel ter
8	123.6	65.7	9264 13 AF014386	AF014386 Drosophila
9	63.4	33.7	190 6 BD173516	BD173516 Novel ter
10	63.4	33.7	190 6 BD177020	BD177020 Novel ter
11	63.4	33.7	8550 13 AF183905	AF183905 Black que
12	46.4	24.7	187 6 BD173512	BD173512 Novel ter
13	46.4	24.7	187 6 BD177016	BD177016 Novel ter
14	46.4	24.7	9275 13 AB017037	AB017037 Himecobi
15	46.4	24.7	9275 13 AB183472	AB183472 Himecobi
16	42.6	22.7	184134 14 AC151155	AC151155 Bos tauru
17	42.6	22.7	220694 14 AC155178	AC155178 Bos tauru
18	40.8	21.7	140570 14 AC132079	AC132079 Felis cat

19	40.8	21.7	141299	14 AC132080	AC132080 Felis cat
20	40.6	21.6	118765	5 BX649275	BX649275 Zebrafish
21	40	21.3	110263	8 HS681N20	HS681N20 Human DNA
22	40	21.3	178377	14 AC165393	AC165393 Pan trogl
23	40	21.3	215706	14 AC161192	AC161192 Pongo pyg
24	39.6	21.1	156604	14 AC069371	AC069371 Homo sapi
25	39.6	21.1	160624	14 AC144418	AC144418 Rattus no
26	39.6	21.1	190904	8 AP002358	AP002358 Homo sapi
27	39.6	21.1	226020	14 AC087689	AC087689 Homo sapi
28	38.6	20.5	221 8 HS197H1R	HS197H1R	Z65084 H. sapiens C
29	38.6	20.5	53711 8 AB038985	AB038985 Homo sapi	AB038985 Homo sapi
30	38.6	20.5	70185 14 AC091168	AC091168 Homo sapi	AC091168 Homo sapi
31	38.6	20.5	179216 14 AC074016	AC074016 Homo sapi	AC074016 Homo sapi
32	38.6	20.5	181150 8 AC007256	AC007256 Homo sapi	AC007256 Homo sapi
33	38.4	20.3	194464 14 AC160270	AC160270 Bos tauru	AF536531 Aphid let
34	38.2	20.3	9812 13 AF536531	AF536531	AF536531 Aphid let
35	38.2	20.3	9829 6 AX763949	AX763949 Sequence	AX763949 Sequence
36	37.8	20.1	226138 14 AC103252	AC103252 Rattus no	AC103252 Rattus no
37	37.8	20.1	228760 14 AC149728	AC149728 Bos tauru	AC149728 Bos tauru
38	37.8	20.1	251750 14 AC097029	AC097029 Rattus no	AC097029 Rattus no
39	37.4	19.9	5504 6 AX251954	AX251954 Sequence	AX251954 Sequence
40	37.4	19.9	5504 6 AX344346	AX344346 Sequence	AX344346 Sequence
41	37.4	19.9	5504 6 AX349003	AX349003 Sequence	AX349003 Sequence
42	37.4	19.9	69248 8 AC113427	AC113427 Homo sapi	AC113427 Homo sapi
43	37.4	19.9	166777 14 AC106813	AC106813 Homo sapi	AC106813 Homo sapi
44	37.4	19.9	168347 14 AC025336	AC025336 Homo sapi	AC025336 Homo sapi
45	37.4	19.9	190024 8 AC122714	AC122714 Homo sapi	AC122714 Homo sapi

ALIGNMENTS

RESULT 1	BD173514	188 bp	RNA	linear	PAT 18-FEB-2003
LOCUS	BD173514	Novel tertiary structure having ability to accelerate translation			
DEFINITION	Novel tertiary structure having ability to accelerate translation				
ACCESSION	BD173514.1	GI:28414845			
KEYWORDS	WO 02061080-A/4.				
SOURCE	Cricket paralysis virus				
ORGANISM	Cricket paralysis virus				
REFERENCE	1 (bases 1 to 188)				
AUTHORS	Nakashima, N. and Kanamori, Y.				
TITLE	Novel tertiary structure having ability to accelerate translation				
JOURNAL	Patent: WO 02061080-A 4 08-AUG-2002;				
	JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF				
	SERICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE				
	FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA, YASUSHI KANAMORI				
COMMENT	OS Cricket paralysis virus				
	PN WO 02061080-A/4				
	PD 08-AUG-2002				
	PF 31-JAN-2001 WO 2001JP000641				
	PR 25-JAN-2001 JP 01P 016746				
	PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI				
	PC C12N15/11, C12N15/86, C12P21/02				
	CC Novel tertiary structure having ability to accelerate translation				
	Novel tertiary structure having ability to accelerate translation				
	translation activity				
	Key source				
	Location/Qualifiers				
	1.188				
	/organism="Cricket paralysis virus"				
	/molecule="genomic RNA"				
	/db_xref="taxon:12136"				
FEATURES	source				
	1.188				
	/organism="Cricket paralysis virus"				
	100.0%; Score 188; DB 6; Length 188;				
	Best local similarity 66.5%; Pred. No. 6.4e-38;				
ORIGIN	Query Match				
	Best local similarity				

Matches 125; Conservative 63; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAAAAAGUGAUCUUCGUAUUAACAUAUUGAGAGUUAUUAUAUACAAGAGUGC 60
Db 1 CAAAAATGATCTGCTTGTGTAATATCAATTTTGAGAGTTAATAATTAACAAGTAGTC 60

Qy 61 UAUUUUUUAUUUAGUUAUGCUAUUUUACUUUUCUCCAGAGUCCUAGUGGACGCC 120
Db 61 TATTTTGTATTTAGTTAGCTATTACTTACGTTCCAGAGATGCCATGAGGACGCC 120

Qy 121 ACAUAUUCAGAGAGACCCUUCUGCGUUUUUACAGUUAUGUAGUGCAAAAACCUAAGAA 180
Db 121 ACAATATCCAGAGAGCCCTCTGCGGTTTTTCAGATTAGTAGTGCAGAAAACTAAGAA 180

Qy 181 AUUUAACU 188
Db 181 ATTACCT 188

RESULT 2
BD177018

LOCUS BD177018 188 bp RNA linear PAT 16-APR-2003

DEFINITION Novel translational activity-promoting higher-order structure.

ACCESSION BD177018

VERSION BD177018.1 GI:30014278

KEYWORDS JP 2002306168-A/4.

SOURCE Cricket paralysis virus

ORGANISM Cricket paralysis virus

Viruses; ssRNA positive-strand viruses, no DNA stage;
Dicistroviridae; Crispavirus.

REFERENCE 1 (bases 1 to 188)
Nakashima, N. and Kanamori, Y.
Novel translational activty-promoting higher-order structure
Parent: JP 2002306168-A 4 22-OCT-2002;
DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND
ENTOMOLOGICAL HIROSHIMA PREFECTURAL INDUSTRIAL TECHNOLOGY PROMOTION
ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND
FISHERIES

FEATURES

OS Cricket paralysis virus

PN JP 2002306168-A/4

PD 22-OCT-2002

PF 25-JAN-2001 JP 2001016746

PI NOBHIIKO NAKASHIMA, YASUSHI KANAMORI

PC C12N15/09, C12N15/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02// PC
(C12N15/09, C12R1:92), C12N15/00, C12N5/00, (C12N15/00, C12R1:92) CC
Novel translational activty-promoting higher-order structure FH

Key Location/Qualifiers

FT source 1..188
/organism="Cricket paralysis virus".
location/Qualifiers
1..188
/organism="Cricket paralysis virus"
/mol_type="genomic RNA"
/db_xref="taxon:12136"

FEATURES

source

1..188
/organism="Cricket paralysis virus"
/mol_type="genomic RNA"
/db_xref="taxon:12136"

ORIGIN

Query Match 100.0%; Score 188; DB 6; Length 188;
Best Local Similarity 66.5%; Pred. No. 6,4e-38;
Matches 125; Conservative 63; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAAAAAGUGAUCUUCGUAUUAACAUAUUGAGAGUUAUUAUAUACAAGAGUGC 60
Db 1 CAAAAATGATCTGCTTGTGTAATATCAATTTTGAGAGTTAATAATTAACAAGTAGTC 60

Qy 61 UAUUUUUUAUUUAGUUAUGCUAUUUUACUUUUCUCCAGAGUCCUAGUGGACGCC 120
Db 61 TATTTTGTATTTAGTTAGCTATTACTTACGTTCCAGAGATGCCATGAGGACGCC 120

Qy 121 ACAUAUUCAGAGAGACCCUUCUGCGUUUUUACAGUUAUGUAGUGCAAAAACCUAAGAA 180
Db 121 ACAATATCCAGAGAGCCCTCTGCGGTTTTTCAGATTAGTAGTGCAGAAAACTAAGAA 180

Qy 181 AUUUAACU 188

Db 181 ATTACCT 188

RESULT 3
AF218039

LOCUS AF218039 9185 bp RNA linear VRL 02-JUL-2000

DEFINITION Cricket paralysis virus nonstructural polyprotein and structural polyprotein genes, complete cds.

ACCESSION AF218039

VERSION AF218039.1 GI:8895506

KEYWORDS

SOURCE AF218039.1 GI:8895506

REFERENCE 1 (bases 1 to 9185)
Cricket paralysis virus
Viruses; ssRNA positive-strand viruses, no DNA stage;
Dicistroviridae; Crispavirus.

REFERENCE 2 (bases 1 to 9185)
Wilson, J.E., Powell, M.J., Hoover, S.E. and Sarnow, P.
Naturally occurring dicistronic cricket paralysis virus RNA is
regulated by two internal ribosome entry sites
Mol. Cell. Biol. 20 (14), 4990-4999 (2000)

JOURNAL 10866656

REFERENCE 2 (bases 1 to 9185)
Wilson, J.E., Powell, M.J., Hoover, S.E. and Sarnow, P.
Submitted (20-DEC-1999) Microbiology & Immunology, Stanford
University, 299 Campus Drive, Stanford, CA 94305, USA

FEATURES

source

Location/Qualifiers

1..9185
/organism="Cricket paralysis virus"
/vrlion
/mol_type="genomic RNA"
/db_xref="taxon:12136"
709..6024
/codon_start=1
/product="nonstructural polyprotein"
/protein_id="AAF80998.1"
/db_xref="GI:8895507"
/translation="MSFQQTNNATNNINSLEELAQELIAAQFSGNDLDFCTRYQ
SKPQLDLSBECYCDNDPDCDRIKREBELKLIPLTSDVGVNFEKKGILVMKRVQ
NYCEPRVYSTFNGGLIVSPRFMDHLDWFQOMKLVSNDECRALFKRKTLLMSGD
VESNRPVQSRPVACDNDPRAIRLEKALQREDEKISTLIKLRQBIKNNRYTGQFF
DLKAKGEGQLNGNLFRICDFLENSLPETLAQIOTVLTDTDKYVNLKEDLVYAI
LLVLRLVLMWKKRYAALVILVLFVWHFPGPKQIOTVLTDLKDKILTOTTQAGTELT
EEVYVHPMDQKQILPVALFAFPAIKIKPGQDMNNYISRLDIRIKATIGSKTYDVC
SEYFNLSDVEVKYVLGKELKTQGLYDEHFMWAGIRHYLDLDRNKTLTDETAAK
VDLYRGLKLYSEERKIPBDIARITTYMLFPKSLYEVLSPVGGGPKMPTVWL
TGESIGIKTQMIYPLCIDILREMGIVKPDVAGQAVARQVETEVGDYNGQKIVYDD
AFQLDDDKTKEPPIFEVLRICNTPEPHLMAALQDKMYSAEVLATYTNQVQLE
SITFPDAFYNNRKTAAVRYQIKQEKSIWRNARGENVALDVTKLKDEADLSYVEF
OKMRFPDSATKTIWIDGEPISYDERARTICAMKEKEKTHQLOMLRAVSRTVAOC
GSETEYVDVMDETYFNLISGPMAGSLIEMEPASDATPNAIYIEKNIKET
KSKKMTITLDEOISALSTIKRLKNAKAFISEHHYLTALGTAGVAWISAFMYSFER
KLTKEVSGSGDKTKQIKSRVAVEGKELEKAEATQSGSDPAHAALVLDLQKNTYCL
KTYKAVGEMKRYRLATAPLFGVNCMPKPYHIEFLYARKVAPSTNIYFSGNCDVI
VVPVSHPIAPNAERVELTACTRIHYKDETPDCVLNLMHRMKPHRDLKHFYKS
DOGNRFGPGTLAFHOSANETICRAYOMLAIRLDQETIRIYHEDTDMFDESSYTS
QKDCYENAPTYTGNGSITVGLYNKRMERKLGKMT PGNVSCHEGYACPLOTALMDG
LNRLEKLPVNNITVQCFFEPSDIKIDTMSGETPBGKCAIGKSIKVGOAVKTLTLK
SCITYMLSKPIPAHLNTRLPNGEIVDPLMKGKKGCVDTAVLADIVESAALDVLK
QVLLIQVNSMLDNVYKRFLLYEBAITQGTGDDFMKGIARQTS PGARYFQMRKLPKG
ODMMSGGOYDPTSGRAOELRDBVELDINCAKGIKIVVAFVDTLKDRRPLEKVDAG
KTRVPSAGBQHVVAFRKYPLPFAVYLMNNITDNEIATGVTVSTDMERLAKRLKKG
NKVINGADGNPDGSLVAPFPGSGCKSYPEFKTINDVNTDGRKNLITCGLKTHIV
HSVHSIGNVNTMTHSOPSNGPFTVINCILNYSMTWRVWLLAKLAPKQSMKFR
ENVSMISYGDNCLNISDRVVEWFOQITISQMEIKHEYDBGTGNTGVAFPSISEI
HFLKRFVFSHQLORTVAPLQKDVLYEMLNTYNTIDNEILMNNINTAFREIYVHGK
SEYQCLRSGIEDLWKGILPQOPQILTRFKAYLMDATMLADEYDE"

/codon_start=1
/product="structural polyprotein"
/protein_id="AAF80999.1"

CDS

CDS

[illegible]

ORIGIN

100.0%; Score 188; DB 13; Length 9185;

Matches 125; Conservative 63; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAATGUGAUCUCUCUGUAAUACAAUUTUUGAGAGUUAUAUAUUAACAAGUAGUC 60
Db 6029 CAAAATGATCTTGGCTTGAATACAAATTTTGAGAGGTTAAATTAACAAGTAGTGC 6088

QY 61 UAUUUUUGUAUUAGGUAACUUUUAGCUUUAGCUCUCCAGAGGCCUAUGSGCAAGCCCC 120
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::
Db 6089 TATTTTGTATTGGTTAGCTAATTTTAGCTTTTCAGTTCAGAAATCCCTAATGGCGAGCCCC 6148
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::

OY 121 ACAAUUCCAGGAA GCCCUCUCUGCGGUUUUUCAGAUTUAGUGACGAAAAA CCUAAGA 180

D6 6149 ACATATCCAGGAAGCCCTCTCTGTGGTATTTCAGATTAGGTAGCGAAAAA CCTTAGA 6208

OY	181	AUUUACCU	188
	:::		
DB	6209	ATTACCT	6216

RESULT 4	
AR496112	
LOCUS	1345 bp DNA
DEFINITION	Sequence 1072 from patent US 6703491.
	linear PAT 22-SEP-2004

ACCESSION	AR496112	GI:52431587
VERSION	AR496112.1	
KEYWORDS		
SOURCE	Unknown.	

ORGANISM	REFERENCE
Unknown.	1 (bases 1 to 1345)
Unclassified.	Homburger, S. A., Ebens, A. J. Jr., Erickson, C. S., Francis-Lang, H. L.,
	Marzolis, J. S., Reddy, B. P., Ruddy, D. A. and Buchanan, A. R.

FEATURES	TITLE
Exelxits, Inc., South San Francisco, CA	Drosophila sequences
Location/Qualifiers	Patent: US 6703491-A 1072 09-MAR-2004;

```
source      1. .1345
            /organism="unknown"
            /mol_type="genomic DNA"
```

Query Match	Score 125.2;	DB 6;	Length 1345;
Best Local Similarity	51.1%	Pred. No. 7e-22;	
Matches 95;	Conservative 53;	Mismatches 38;	Indels 0;
			Gaps 0;

Oy 3 AAAAUGCAGUCGUGCUGUAAAUAACAUUUUUGAGAAGCUAAUAAAUTUAACAAGUAGUCUA 62
|||:::||:::||::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 520 AAGAATCATCTTGGTCCTTAATACAATTGTGAAGGGTTAATTAACAGGAAGTAGTGCTA 579

Dy 63 TTTTGTGTAATTTAGAGCGTAACGTAATTTTAGAGCTTTAAGCCGTAAGGCCGAGCCCCAAC 122
::: : ::|||:: ||::|| |::||| : |||||::| : |||||
Db 580 TTTTAAATTATTAGGGTTAACTATTGATTTCATCTGTTACAGGAATGCTTATTGGCAGGCCCAT 639

QY 123 AAUATCCAGAAAGCCCTUCUCGCGUUUUUCAGAUUAGUUGCGAAAAACCUAAGAAU 182

Db 640 AATATCCAGACACCCCTCTCTCTTATATGATATAGTTCATTTAGATAAGAA 699

Qy	183	UVACCU	188
	:		:
Db	700	TAACCT	705

RESULT 5

ARS11394

DEFINITION	Sequence	Accession	16354	From patent	US 6703491.
ACCESSION	AKS11394				
VERSION	AKS11394.1	GI:52446869			

KEYWORDS
SOURCE
ORGANISM
UNKNOWN.
UNKNOWN.
UNKNOWN.
UNCLASSIFIED.
1 (BASE 1 TO 1145)

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (pages 1 to 1937)	Homburger, S.A., Ebers, A.J. Jr., Erickson, C.S., Francis-Lang, H.L., Margolis, J.S., Reddy, B.P., Ruddy, D.A. and Buchman, A.R.	Drosophila sequences	Patent: US 6703491-A 16354 09-MAR-2004

```

FEATURES
  source
    Exelixis, Inc.; South San Francisco, CA
    location/Qualifiers
      1. 1345
      /organism="unknown"

```

ORIGIN	/mol_type="genomic DNA"
Query Match	66.6%; Score 125.2; DB 6; Length 1345;
Best local similarity	51.1%; Pred NO. 79-22.

3' AAAATGTAUUCUUGCUUGAAUACAAUUTUGAGAGUAAUAAUUAUACAAGACGCUA 62
 :::::::::::::::::::::
 95; Conservative 53; Mismatches 38; Indels 0; Gaps 0

Db 520 AAGATGATATCTTGGCTCTCTTATACAA TTTTGAAGAGTTAAATAAAGAAGGAAGTATGCTA 579

[illegible]

QY	183	UWACCU	188
	:	:	
Db	700	TAACCT	705

RESULT 6
BD173513
LOCUS
DN17771.1 100 bp
DNN
149897 DAT 18-FEB-2000

LOCUS	BD173513	107 bp	348	aligned	50 to 1000
DEFINITION	Novel tertiary structure having ability to accelerate translation activity.				
ACCESSION	BD173513				
VERSION	BD173513.1	GI:28414844			

KEYWORDS	MO 02061080-A/3.
SOURCE	Drosophila C virus
ORGANISM	Drosophila C virus
	viruses; ssRNA positive-strand viruses, no DNA stage;

REFERENCE
AUTHORS
TITLE

Dicistroviridae; Crispavirus.
1 (bases 1 to 189)
Nakashima, N. and Kanamori, Y.
Novel tertiary structure having ability to accelerate translation
activation

JOURNAL
Patent: WO 02061080-A 3 08-AUG-2002;
JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF
SERICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE
FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA, YASUSHI KANAMORI

COMMENT	OS Drosophila C virus PN MO 02061080-A/3 ED 08-AUG-2002 PF 31-JAN-2001 WO 2001JP000641 PR 25-JAN-2001 JP 01P 016746 PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI PC C12N15/11, C12N15/86, C12P21/02 CC Novel tertiary structure having ability to accelerate translation activity
FEATURES	FH Key Location/Qualifiers FT source 1..189 1..189 /organism='Drosophila C virus'. Location/Qualifiers 1..189 /organism='Drosophila C virus' /mol_type='genomic RNA' /db_xref='taxon:64279'
ORIGIN	Query Match 65.7%; Score 123.6; DB 6; Length 189; Best Local Similarity 51.1%; Pred. No. 2.5e-21; Matches 95; Conservative 52; Mismatches 39; Indels 0; Gaps 0;
Qy	3 AAATGUGAGUCGUCGUGAAUUAACAUUUUGAGAGUAAUUAUAACAAGUGCCUA 62
Db	4 AAGATGTGATCTGGCTCCTTATACATTTTGAGAGGTTAATTAAGAAGAGTATGCTA 63
Qy	63 UUUUUGAUNUUGAUUAGUUUUUAGUUUUGUUCGAGAGUCCUAGUGCCAGCCAC 122
Db	64 TTTTAATTAATTAGTTAATCTATTGTTTACTGTTCAGAGATGCTATTGCGAGCCCAT 123
Qy	123 AAUAVCCAGAGAGCCUUCUGCGGUGUUUUCAGAUUAGUAGUGCAAAACCUAAGAAU 182
Db	124 AATATCCAGAGACCCCTCTCTCTTATATGATTAGGTTGTCTTTAGATTAAGAAA 183
Qy	183 UUACCU 188
Db	184 TAACT 189
RESULT 7	
BD177017	
LOCUS	BD177017 189 bp RNA linear PAT 16-APR-2003
DEFINITION	Novel translational activity-Promoting higher-order structure.
ACCESSION	BD177017
VERSION	BD177017.1 GI:30014277
KEYWORDS	JP 2002306168-A/3.
SOURCE	Drosophila C virus
ORGANISM	Drosophila C virus Virusess; ssRNA positive-strand virusess, no DNA stage; Dicistroviridae; Crispavirus. 1 (bases 1 to 189) Nakashima,N. and Kanamori,Y. Novel translational activity-promoting higher-order structure Patent: JP 2002306168-A 3 22-OCT-2002; DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND ENTOMOLOGICAL HIROSHIMA PREFECTURAL INDUSTRIAL TECHNOLOGY PROMOTION ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND FISHERIES
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	OS Drosophila C virus PN JP 2002306168-A/3 PD 22-OCT-2002 PF 25-JAN-2001 JP 2001016746 PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI PC C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02// PC (C12N15/09, C12R1/92), C12N15/00, C12N5/00, (C12N15/00, C12R1/92) CC Novel translational activity-promoting higher-order structure FH Key Location/Qualifiers FT source 1..189 1..189 /organism='Drosophila C virus'. Location/Qualifiers 1..189 /organism='Drosophila C virus' /mol_type='genomic RNA'
FEATURES	
SOURCE	

[illegible]

CDS

NSBYGDCIDFVDIADKRLQKGTLEIEFDYASDPENFTQYHFKSTIKPASMOKYK
 DMDICSDCKTYLAKYBEIKTILAGMGLVSLALAMYMYSKSLDPVE
 AEVAPSGDATVRLPRKLVIEGASGVDTOKIVPVUTETHEHRNKEIETSCBCEM
 HRNSAFNMMDTEEDNCTIEDLNQKRELBQNSKTQSWLDRKRCVITAB
 VGGSDVKTNAQIKRVEAGABELVYALTOGCSDDAHMIDVFOKNTYRMSYFG
 DKRYQDNCSTFVRGMSFTMPYHFOVAFARLPNTTISIQMSSELMQPLSHFS
 AGVNPFLYTDNCVLPFGNVDLPCNNVLTNSRCKTDRHDLRHFIITSDQKSGF
 SGAAATPHNNMGLYRYVNMNAYRPCDKLEIHPHDPGFPESTYIORDCYVNP
 TRDDCCSIIICLYNRYLKERKIGHNANGNAEENGYACPLTDELETRAFSLYNNK
 NISQPYEIPNMYDPLGDSVPEGKTFALGKSVIGQAVNSITISRTKLSVFT
 MKPPLKPTTILNNRVNPLSLGKCKGVDAVLSDDEVLSAQOVCEVLMNQYNLN
 KTKYQRLTYEELRGTODEFMCAINRTTSPGPYAKMKNAGKQOMMGFGSEFP
 TSNVALARKVEOLIEDCASGISNYLFDVTLKDBRDIAKVNQKTVSAGPQF
 VVAROYELPRAMLMENRISNEVAVGTVYSDMERIARLTKGSHVLADGNPD
 GSLVAQILMAIFWEITPVYMLKQFIDINSEBKRILCTCLGMSLHYSHVYENYTM
 WTHSQPSGNPTVILINCLYNSIIMLSHWRMEKFOPLSKMKNFNSYVALITYGDN
 VLINIDAVKEVFNQINISEVWTEMRHEVTEAKTGDIYKSKEDIFPLRKRFSPB
 LORVAPLKEIYIEMLNMSRSDIDPEILMNSIETAFREVVYHGKEEYDKLRSAVLA
 LKVOELPENQILTYNQYLHDIYADPLVDP"
 <6267..8972
 /note="ORF-2; putative; translation start site unknown"
 /codon_start=1
 /evidence=experimental
 /product="capsid polyprotein"
 /protein_id="AAC58808.1"
 /db_xref="GI:2388674"

/translation="ANFOTNNNNIENEDRKITSOKETVHRSRGVPTTAVDYS
 LSTYLSMTTREDRIHTIKDPLSRPILITGLMSATTAEQLTANPEVFSINTY
 QEKLRGFGVLRATLVIVOVNSOPFOQGRMLQYPAQVMPNRSVLSNLTQSRSC
 PRDLDSVGEVEMRIYVYSVHYNLITQGSFGALYLVVYSQLRDVGTSEVY
 TVMHLSDVDVQYPTGANIFITGSSPNFASIKQDKDEKLDLWTSKAYAKOD
 KIPNOVASEITOLKSGTISGIGVSEGLSTMKITILINMFKPMIAVONIRK
 MGRSKRTVQGLPESBKRGQRYRANPDGADTSIKALASONEIETSGLSGSPDM
 DLSVLSIPNEMDFITNNTDASISLIDNVTIKPYSSTILDPKRCITGVYAN
 THGWCSIIYVTFKRVKTOFSGRLISFLPYNTTISAGVPVSTOKIVDLKTS
 TEVSFTIPYVSFPMYCIIRPEASMLGTDMANAVTISAVRVSLNQLVANNVPS
 IDITVEGSGDLTPAAPMAPSVYVSGGFTLADDAAKAKORSEYNNIIPORTSNG
 KREVEDARIVAGNEDLATORNDAGHVPMTIDTKHIDSNMPEAHGCEKMSIR
 OLIRFEMALNSLNLISDAPTLTAPSVQHPPTVYAPABMSFEYIYFISGTWRG
 MRFLOAVRTNSAETSVKTDITVTNLNMSVDSFNSLINFSTIDYPIKSTGLMPG
 TSGFNSMTYIDPEVEGFMEFLPYNISHISPATYVRGTESEITITNSVLRGLHPQ
 IYVAPPGTITATDVVNAOPARAPSDFSFWYLVGPPLTVNAP"

ORIGIN

Query Match 65.7%; Score 123.6; DB 13; Length 9264;

Best Local Similarity 51.1%; Pred. No. 1.3e-21;

Matches 95; Conservative 52; Mismatches 39; Indels 0; Gaps 0;

QY 3 AAAAUGGAGUCUCUGUAAUACAUAUUUGAGAGGUAAUAAUUAACAAGUGUGUA 62
 DB 6081 AAGATGATGATCTTCTCTTAAATTTGAGAGGTTAATTAAGAAAGATGATGCTA 6140
 QY 63 UUUUUUUAUUAGUGUACUUAUUGCUUUGAGUGGCUAGUGGCCCCAC 122
 DB 6141 TCTTAATTAATTAAGTTAATTAAGTTTACGTTCAGAGATGCTATTTGCAAGCCCAT 6200
 QY 123 AAUAUCCAGAGAGCCUCUCUGGCUUUUUUUAUUGUGUGUAGAAAACUAGAUAU 182
 DB 6201 AATATCCAGAGACCCCTCTGCTTCTTATATGATTAGTTGATTTAGAAAGAAA 6260
 QY 183 UUAACU 188
 DB 6261 TAACTT 6266

RESULT 9 BD173516 190 bp RNA linear PAT 18-FEB-2003
 LOCUS BD173516
 DEFINITION Novel tertiary structure having ability to accelerate translation
 activity.
 ACCESSION BD173516
 VERSION BD173516.1 GI:28414847
 KEYWORDS WO 02061080-A/6.

SOURCE unidentified

ORGANISM unidentified

REFERENCE unclassified

AUTHORS 1 (bases 1 to 190)

TITLE Nakashima, N. and Kanamori, Y.

JOURNAL Novel tertiary structure having ability to accelerate translation activity

Patent: WO 02061080-A 6 08-AUG-2002;

JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF

SERICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE

FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA, YASUSHI KANAMORI

OS Black queen-cell virus

PN WO 02061080-A/6

PD 08-AUG-2002 NO 2001JP000641

PR 31-JAN-2001 JP 01P 016746

PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI

PC C12N15/11, C12N15/86, C12P21/02

CC Novel tertiary structure having ability to accelerate CC

translation activity

FH Key Location/Qualifiers

FT source 1..190

Location/Qualifiers

1..190

/organism="Black queen-cell virus"

/mo_type="genomic RNA"

/db_xref="taxon:32644"

ORIGIN

Query Match 33.7%; Score 63.4; DB 6; Length 190;

Best Local Similarity 42.7%; Pred. No. 7.9e-06;

Matches 79; Conservative 43; Mismatches 61; Indels 2; Gaps 2;

QY 1 CAAAAGUGAGUCUCUGUAAUAC-AAUUUGAGAGGUAAUAAUUAACAAGUGUG 59
 DB 2 CAACAATGATATCTTCTGCGAGGCAAAATTTGCACAGTAAATCTGCACAGTAGTG 61
 QY 60 CUUUUUUUAUUAGUGUACUUAUUGCUUUGAGUGGCUAGUGGCCCC 118
 DB 62 CTATGTGGAATCACCGTACTTATTAAGTTACGCTCAAGATCGGTGATGCGAGCC 121
 QY 119 CCACAUAUCCAGAGCCUCUCUGGCUUUUUUUAUUGAGUGUGUAGUAGUAGUAG 178
 DB 122 CTATCATATCTAGAGAACTGTGCTATGTTAGAGATTAAGTAGTCTCTAAACAGAAC 181
 QY 179 AAUUU 183
 DB 182 AATT 186

RESULT 10

BD177020 190 bp RNA linear PAT 16-APR-2003

LOCUS BD177020

DEFINITION Novel translational activity-promoting higher-order structure.

ACCESSION BD177020

VERSION BD177020.1 GI:30014280

KEYWORDS JP 2002306168-A/6.

SOURCE unidentified

ORGANISM unidentified

REFERENCE 1 (bases 1 to 190)

Nakashima, N. and Kanamori, Y.

Novel translational activity-promoting higher-order structure

Patent: JP 2002306168-A 6 22-OCT-2002;

DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND

ENTOMOLOGICAL, HIROSHIMA PREFECTURAL INDUSTRIAL TECHNOLOGY PROMOTION

ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND

FISHERIES

OS Black queen-cell virus

PN JP 2002306168-A/6

PD 22-OCT-2002

PF 25-JAN-2001 JP 2001016746

	PI	NOBUHIKO NAKASHIMA, YASUSHI KANAOKORI PC C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02// CC (C12N15/09, C12R1.92), C12N15/00, C12N5/00, (C12N15/00, C12R1.92) CC Novel translational activity-promoting higher-order structure FH
ORIGIN	Key	Location/Qualifiers
	FT source	1..190
		/organism='Black queen-cell virus'.
FEATURES	source	Location/Qualifiers
		1..190
		/organism="unidentified"
		/mol_type="genomic RNA"
		/db_xref="taxon:32644"
Query Match	33.7%; Score 63.4; DB 6;	Length 190;
Best Local Similarity	42.7%; Pred. No. 7.9e-06;	
Matches	Conservative 43; Mismatches 61; Indels 2; Gaps 2	
OY	1	CAAAAGUCGACUCCUGCUAUAUAC-AAUUUGAGAAGUUAAUUAUACAUAUG 59
Dd	2	CACAATGTGAATCTTGGCTTGCGAGGCAAAATTTCACAGTAATAAACCTGCMAACTAGT 61
OY	60	CUAUUUUGUUUUUAGUAGUAGUUUAGUUUAGUUUCAAGUCCAG-UGCCAGCC 118
Dd	62	CTATTGTGGAAATCACCGTACCCTATTTAGCTTACGCTCCAAGTCGGTGATACAGCC 121
OY	119	CCACAUAUCCAGGAAGCCCUCUCGCGUUUUUUCAGAUAGUUAGUAGUAAAAACCUAAG 178
Dd	122	CTATCATATATCTAGAGAACTGCTAGTTAGTAAGAAATTAAGTATGTCCTATAACAGAAC 181
OY	179	A AAUU 183
Dd	182	A AATT 186
RESULT 11		
LOCUS	AF183905	8550 bp ss-RNA linear VRL 16-AUG-2000
DEFINITION	Black queen cell virus nonstructural polypeptide (orf1) and structural polypeptide (orf2) genes, complete cds.	
ACCESSION	AF183905	
VERSION	AF183905.1	GI:8100530
KEYWORDS	.	
SOURCE	Black queen cell virus	
ORGANISM	Black queen cell virus	
	Viruses; ssRNA positive-strand viruses, no DNA stage; Dicistroviridae; Citrivirus. 1. (bases 1 to 8550) Leat,N., Bail,B., Govan,V. and Davison,S. Analysis of the complete genome sequence of black queen-cell virus, a picorna-like virus of honey bees J Gen. Virol. 81 (Pt 8), 2111-2119 (2000)	
JOURNAL	PUBMED	10900051
REFERENCE	Leat,N. and Davison,S. Direct Submission Submitted (06-SEP-1999) Microbiology, University of Western Cape, Modderdam Rd., Cape Town, Western Cape 7535, South Africa	
TITLE	Journal	Location/Qualifiers
FEATURES	source	1..8550
		/organism="Black queen cell virus"
		/mol_type="genomic RNA"
		/strain="South African"
		/specific_host="Apis mellifera"
		/db_xref="taxon:92395"
		658..5625
		/gene="Orf1"
		658..5625
		/codon_start=1
		/product="nonstructural polypeptide"
		/protein_id="AAF2337.1"
		/db_xref="GI:8100531"
gene		
CDS		

[illegible]

Db 5768 CTACCAATATCTAGAGAACTGTCTATGTTTGAAGATTGCTGTCTTAACAGAAC 5827
OY 179 AAUUV 183
Db 5828 AATT 5832

RESULT 12
BD173512
LOCUS BD173512
DEFINITION Novel tertiary structure having ability to accelerate translation activity.
ACCESSION BD173512.1 GI:28414843
VERSION WO 02061080-A/2.
KEYWORDS Himetobi P virus
SOURCE Himetobi P virus
ORGANISM Dicteroviridae; Crispavirus.
REFERENCE 1 (bases 1 to 187)
AUTHORS Nakashima,N. and Kanamori,Y.
TITLE Novel tertiary structure having ability to accelerate translation activity
JOURNAL JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA,YASUSHI KANAMORI
COMMENT OS Himetobi P virus
PN WO 02061080-A/2
PD 08-AUG-2002
PR 31-JAN-2001 WO 2001JP000641
PR 25-JAN-2001 JP 01P 016746
PI NOBUHIKO NAKASHIMA,YASUSHI KANAMORI
PC C12N15/11,C12N15/86,C12P21/02
CC Novel tertiary structure having ability to accelerate CC translation activity
FH Key Location/Qualifiers
FT source 1..187
Location/Qualifiers
1..187 /organism="Himetobi P virus".
/mol_type="genomic RNA"
/db_xref="taxon:81583"

FEATURES
source
1..187 Location/Qualifiers
/organism="Himetobi P virus"
/mol_type="genomic RNA"
/db_xref="taxon:81583"

ORIGIN
Query Match 24.7%; Score 46.4; DB 6; Length 187;
Best Local Similarity 36.6%; Pred. No. 0.19;
Matches 63; Conservative 37; Mismatches 71; Indels 1; Gaps 1;

OY 15 UGCUUGUAAUACAUAUUUGAGAGUUAUAUUAUACAUGUGCUAUUUUGUUAUUA 74
Db 15 TGATTAGAAAGTAAATAATCTCTAGTTATATATTTTAATGCTCATTTTAAAGAC 74
OY 75 GGUUAGCUAUUAGCUUUAUGUUGCAGAUUGCCUAGUGCGACCCCAAAUUAUCCAGGAA 134
Db 75 CCTTAGTATTATTAACCTTACCGCCAGAGATGGGTGCAGCGCTTCTGCAATATCCAGGCG 134
OY 135 GCCCUCUCUGCGUUUUUACAUGUAGUGAAGGAAAAACUAGAUAUUUAC 186
Db 135 A-CCTAGGTGAGCGCTGTATTTAGTGACCTTAGGCTTAAGAATTTTCAC 185

RESULT 13
BD177016
LOCUS BD177016
DEFINITION Novel translational activity-promoting higher-order structure.
ACCESSION BD177016.1 GI:30014276
VERSION JP 2002306168-A/2.
KEYWORDS Himetobi P virus
SOURCE Himetobi P virus
ORGANISM

Viruses; ssRNA positive-strand viruses, no DNA stage;
Dicteroviridae; Crispaviruses.
REFERENCE 1 (bases 1 to 187)
AUTHORS Nakashima,N. and Kanamori,Y.
TITLE Novel translational activity-promoting higher-order structure
JOURNAL Patent: JP 2002306168-A 2 22-OCT-2002;
DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND FISHERIES
COMMENT OS Himetobi P virus
PN JP 2002306168-A/2
PD 22-OCT-2002
PR 25-JAN-2001 JP 2001016746
PI NOBUHIKO NAKASHIMA,YASUSHI KANAMORI
PC C12N15/09,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P21/02// PC (C12N15/09,C12R1:92),C12N15/00,C12N5/00,C12R1:92) CC Novel translational activity-promoting higher-order structure FH Key Location/Qualifiers
FT source 1..187
Location/Qualifiers
1..187 /organism="Himetobi P virus".
/mol_type="genomic RNA"
/db_xref="taxon:81583"

FEATURES
source
1..187 Location/Qualifiers
/organism="Himetobi P virus"
/mol_type="genomic RNA"
/db_xref="taxon:81583"

ORIGIN
Query Match 24.7%; Score 46.4; DB 6; Length 187;
Best Local Similarity 36.6%; Pred. No. 0.19;
Matches 63; Conservative 37; Mismatches 71; Indels 1; Gaps 1;

OY 15 UGCUUGUAAUACAUAUUUGAGAGUUAUAUUAUACAUGUGCUAUUUUGUUAUUA 74
Db 15 TGATTAGAAAGTAAATAATCTCTAGTTATATTTTAATGCTCATTTTAAAGAC 74
OY 75 GGUUAGCUAUUAGCUUUAUGUUGCAGAUUGCCUAGUGCGACCCCAAAUUAUCCAGGAA 134
Db 75 CCTTAGTATTATTAACCTTACCGCCAGAGATGGGTGCAGCGCTTCTGCAATATCCAGGCG 134
OY 135 GCCCUCUCUGCGUUUUUACAUGUAGUGAAGGAAAAACUAGAUAUUUAC 186
Db 135 A-CCTAGGTGAGCGCTGTATTTAGTGACCTTAGGCTTAAGAATTTTCAC 185

RESULT 14
AB017037
LOCUS AB017037
DEFINITION Himetobi P virus genomic RNA, complete sequence.
ACCESSION AB017037
VERSION AB017037.1 GI:3493357
KEYWORDS nonstructural protein precursor; capsid protein precursor.
SOURCE Himetobi P virus
ORGANISM Himetobi P virus
REFERENCE 1 (sites)
AUTHORS Nakashima,N., Sasaki,Y. and Toriyama,S.
TITLE Determining the nucleotide sequence and capsid-coding region of Himetobi P virus: a member of a novel group of RNA viruses that infect insects
JOURNAL Arch. Virol. 144 (10), 2051-2058 (1999)
PUBMED 10550677
REFERENCE 2 (bases 1 to 9275)
AUTHORS Nakashima,N. and Sasaki,Y.
TITLE Direct Submision
JOURNAL Submitted (20-AUG-1998) Nobuhiko Nakashima, National Institute of Sericultural and Entomological Science, Department of Insect Physiology and Behavior, 1-2 Owasahi, Tsukuba, Ibaraki 305-8634, Japan (E-mail:nakajiemisee.affrc.go.jp, tel:81-298-38-6109, Fax:81-298-38-6028)
FEATURES
source
1..9275 Location/Qualifiers

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 04:43:02 ; Search time 271.684 Seconds
(without alignments)
4611.840 Million cell updates/sec

Title: US-10-088-750C-4
Perfect score: 188
Sequence: 1 caaaagugaucugcug.....aaacuaagaauuaccu 188

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq 21:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003as:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004as:*
- 13: Geneseq2004bs:*
- 14: Geneseq2005s:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	188	100.0	200	8	AB280710 Cricket p
2	188	100.0	200	6	AA150547 Cricket p
3	123.6	65.7	189	8	AB280709 Drosophila
4	123.6	65.7	201	6	AA150546 Drosophila
5	63.4	33.7	190	8	AB280712 Black que
6	63.4	33.7	202	6	AA150549 Black que
7	46.4	24.7	187	8	AB280708 Himetobi
8	46.4	24.7	199	6	AA150545 Himetobi
9	40	21.3	130263	6	ABK83573 Human CDN
10	38.6	20.5	460	9	ACH45871 Human foe
11	38.6	20.5	43938	4	AAK77216 Human imm
12	38.6	20.5	45017	4	AAK77217 Human imm
13	38.6	20.5	84476	12	AD097028 Human can
14	38.2	20.3	9829	8	AA160324 Aphid let
15	37.4	19.9	5504	6	ABK31350 Signal tr
16	37.4	19.9	5504	6	AB170571 Chemical1
17	37.4	19.4	5504	6	AA561255 Human gen
18	36.4	19.4	6440	12	ADMA7607 Polynucle
19	35.8	19.0	2480	10	ADA20684 Rice CDNA

20	35.6	18.9	7312	6	AB133815 Human imm
21	35.6	18.9	7312	6	AB170402 Chemical1
22	35.6	18.9	7312	6	AA561350 Human gen
23	35.2	18.7	117829	12	AD097319 Human can
24	34.6	18.4	188	8	AB280707 Plautia s
25	34.6	18.4	197	6	AA150556 CrPV-like
26	34.6	18.4	200	6	AA150544 Plautia s
27	34.6	18.4	200	6	AA150551 CrPV-like
28	34.6	18.4	430	3	AA235832 Plautia s
29	34.6	18.4	7380	4	AA545361 Chemical1
30	34.6	18.4	7380	4	ABK28196 DNA trans
31	34.6	18.4	8246	6	AB132203 Human imm
32	34.6	18.4	110000	2	AA20248_06 Continuation (7 of
33	34.6	18.4	110000	2	AA20248_07 Continuation (8 of
34	34.2	18.2	6219	6	AB132866 Human imm
35	34.2	18.2	6219	6	AA563324 Chemical1
36	33.8	18.0	1365	8	ACA45063 Prokaryot
37	33.8	18.0	11416	6	AB132119 Human imm
38	33.8	18.0	11416	6	AB170136 Chemical1
39	33.8	18.0	11416	6	AA561064 Human gen
40	33.6	17.9	186	8	AB280711 Triatoma
41	33.6	17.9	198	6	AA150548 Human gen
42	33.6	17.9	6289	8	AB210205 Haematopo
43	33.6	17.9	9289	10	AD584197 Human lym
44	33.6	17.9	56153	4	AA546793 Tumour su
45	33.6	17.9	120239	13	ABD33244 Murine ca

ALIGNMENTS

RESULT 1
AB280710
ID AB280710 standard; RNA; 188 BP.
XX
AC AB280710;
XX
DT 15-OCT-2003 (first entry)
XX
DE Cricket paralysis virus derived pseudoknot sequence.
XX
KW Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;
KW albumen; impurity; higher-order structure; intergenic region; IGR-IRIS;
KW internal ribosome entry site; ss.
XX
OS Cricket paralysis virus.
XX
PN WO2003033719-A1.
XX
PD 24-APR-2003.
XX
PF 08-OCT-2002; 2002WO-JP010447.
XX
PR 17-OCT-2001; 2001JP-00319923.
XX
PA (NAAG-) NAT INST AGROBIOLOGICAL SCI.
XX
PA (NAKE-) WAKENTAKU KK.
XX
PI Nakashima N, Shibuya N, Nishikawa S;
XX
DR WPI; 2003-403230/38.
XX
PT Cell-free protein synthesis means in wheatgerm system to establish
PT overexpression of target gene with base sequence sustaining translation
PT activity and function promotion, for producing useful proteins.
XX
PS Claim 1; Page 33; 39pp; Japanese.
CC The invention relates to a cell-free protein synthesis system derived
CC from wheatgerm where there is substantial exclusion of wheatgerm embryo
CC albumen impurities. The novel system uses a sequence having a higher-
CC order RNA structure that promotes translation activity. The higher-order
CC sequence is preferably a "pseudoknot", especially derived from a range of

CC viruses (AB280707-AB280713). This sequence represents the "pseudoknot"
 CC higher-order sequence from the Cricket paralysis virus. The sequence is
 CC used in a construct which may also include an intergenic region and
 CC internal ribosome entry site (IGR-IRES). The method is applicable in
 CC producing useful proteins

Sequence 188 BP; 56 A; 33 C; 36 G; 0 T; 63 U; 0 Other;

Query Match 100.0%; Score 188; DB 8; Length 188;
 Best Local Similarity 100.0%; Pred. No. 1.2e-46;
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAAGUGAUCUGUCUGUUAUAACAUAUUUGAGAGUUAUAUAUUAACAAGUAGUC 60
 DB 1 CAAAAGUGAUCUGUCUGUUAUAACAUAUUUGAGAGUUAUAUAUUAACAAGUAGUC 60
 QY 61 UAUUUUGUAUUUAGGUAUAGCUUAUUUAGCUUACGUAUUGGCAAGUAGCCAGCCCC 120
 DB 61 UAUUUUGUAUUUAGGUAUAGCUUAUUUAGCUUACGUAUUGGCAAGUAGCCAGCCCC 120
 QY 121 ACAUAUCCAGGAAGCCUCUGCGGUUUUACAGAUUAGUAGUAGAAAAACCUAAGAA 180
 DB 121 ACAUAUCCAGGAAGCCUCUGCGGUUUUACAGAUUAGUAGUAGAAAAACCUAAGAA 180
 QY 181 AUUUACCU 188
 DB 181 AUUUACCU 188

RESULT 2
 AAL50547
 ID AAL50547 standard; RNA; 200 BP.

AC AAL50547;
 XX
 DT 19-DEC-2002 (first entry)

DE Cricket paralysis virus RNA sequence.

XX CrPV-like virus; ss; higher-order structure; drug development;
 KW drug production; translational activity-promoting function;
 KW protein synthesis; structural analysis.

OS Cricket paralysis virus.

PN W0200261080-A1.

PD 08-AUG-2002.

PE 31-JAN-2001; 2001WO-JP000641.

PR 25-JAN-2001; 2001JP-00016746.

PA (NAAG-) NAT INST AGROBIOLOGICAL SCI.

PI Nakashima N, Kanamori Y;

DR WPI; 2002-627482/67.

XX Translational activity-promoting higher-order structure of CrPV-like
 PT viruses for protein translation when suitably initiated, useful in
 PT synthesis of proteins and polypeptides of foreign species for application
 PT in drugs.

XX Claim 1; Fig 1-2; 38pp; Japanese.

XX The invention comprises seven RNA sequences (CrPV-like viruses) which
 CC have a higher-order structure that sustains translational activity-
 CC promoting function. The RNA sequences of the invention are useful in the
 CC synthesis of proteins and polypeptides for application in developing and
 CC producing drugs. The RNA sequences of the invention are also useful in
 CC basic research of protein synthesis and structural analysis by the gene
 CC recombinant technique. The present nucleotide represents a Cricket

CC paralysis virus RNA sequence of the invention
 XX Sequence 200 BP; 60 A; 36 C; 37 G; 0 T; 67 U; 0 Other;

Query Match 100.0%; Score 188; DB 6; Length 200;
 Best Local Similarity 100.0%; Pred. No. 1.2e-46;
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAAGUGAUCUGUCUGUUAUAACAUAUUUGAGAGUUAUAUAUUAACAAGUAGUC 60
 DB 1 CAAAAGUGAUCUGUCUGUUAUAACAUAUUUGAGAGUUAUAUAUUAACAAGUAGUC 60
 QY 61 UAUUUUGUAUUUAGGUAUAGCUUAUUUAGCUUACGUAUUGGCAAGUAGCCAGCCCC 120
 DB 61 UAUUUUGUAUUUAGGUAUAGCUUAUUUAGCUUACGUAUUGGCAAGUAGCCAGCCCC 120
 QY 121 ACAUAUCCAGGAAGCCUCUGCGGUUUUACAGAUUAGUAGUAGAAAAACCUAAGAA 180
 DB 121 ACAUAUCCAGGAAGCCUCUGCGGUUUUACAGAUUAGUAGUAGAAAAACCUAAGAA 180
 QY 181 AUUUACCU 188
 DB 181 AUUUACCU 188

RESULT 3
 AB280709
 ID AB280709 standard; RNA; 189 BP.

AC AB280709;
 XX
 DT 15-OCT-2003 (first entry)

DE Drosophila C virus derived pseudoknot sequence.

XX Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;
 KW albumen; impurity; higher-order structure; intergenic region; IGR-IRES;
 KW internal ribosome entry site; ss.

OS Drosophila C virus.

PN W02003033719-A1.

PD 24-APR-2003.

PE 08-OCT-2002; 2002WO-JP010447.

PR 17-OCT-2001; 2001JP-00319923.

PA (NAAG-) NAT INST AGROBIOLOGICAL SCI.

XX (WAKE-) WAKENYAKU KK.

PI Nakashima N, Shibuya N, Nishikawa S;

DR WPI; 2003-403230/38.

XX Cell-free protein synthesis means in wheatgerm system to establish
 PT overexpression of target gene with base sequence sustaining translation
 PT activity and function promotion, for producing useful proteins.

PS Claim 1; Page 32; 39pp; Japanese.

XX The invention relates to a cell-free protein synthesis system derived
 CC from wheatgerm where there is substantial exclusion of wheatgerm embryo
 CC albumen impurities. The novel system uses a sequence having a higher-
 CC order RNA structure that promotes translation activity. The higher-order
 CC sequence is preferably a "pseudoknot", especially derived from a range of
 CC viruses (AB280707-AB280713). This sequence represents the "pseudoknot"
 CC higher-order sequence from the Drosophila C virus. The sequence is used
 CC in a construct which may also include an intergenic region and internal
 CC ribosome entry site (IGR-IRES). The method is applicable in producing
 CC useful proteins

SO Sequence 189 BP; 57 A; 30 C; 34 G; 0 T; 68 U; 0 Other;

Query Match 65.7%; Score 123.6; DB 8; Length 189;
Best Local Similarity 79.0%; Pred. No. 3e-27;
Matches 147; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 3 AAAAUGAUCUUCUCUGUAUAAUACAUAUUUGAGAGUUAUAAUUAACAAGUGUGCUA 62
DB 4 AAGUGAUGAUCUUCUCUGUAUAAUACAUAUUUGAGAGUUAUAAUUAACAAGUGUGCUA 63

QY 63 UUUUUGAUAUUAGGUAUAGCUUAUAGCUUACGUCUCCAGAUCCGCUAGUGCCAGCCCCAC 122
DB 64 UCUUAUAUAUUAGGUAUAGCUUAUAGCUUACGUCUCCAGAUCCGCUAGUGCCAGCCCCAU 123

QY 123 AAUUAUCCAGAGACCCUCUCUGCGUUUUUUCAGAUUAGUGUGGAAAAACCUAAGAAU 182
DB 124 AAUUAUCCAGAGACCCUCUCUGCGUUUAUAGUUAUAGUGUGUUAUAGAAUAAAGAAA 183

QY 183 UUAACCU 188
DB 184 UUAACCU 189

RESULT 4
AAL50546
ID AAL50546 standard; RNA; 201 BP.

AC AAL50546;
XX
DT 19-DEC-2002 (first entry)
DE Drosophila C virus RNA sequence.

XX
KM CrPV-like virus; ss; higher-order structure; drug development;
KW drug production; translational activity-promoting function;
KM protein synthesis; structural analysis.

XX
OS Drosophila C virus.
XX
PN W0200261080-A1.

XX
PD 08-AUG-2002.
XX
PF 31-JAN-2001; 2001WO-JP000641.

XX
PR 25-JAN-2001; 2001JP-00016746.
XX
PA (NAG-) NAT INST AGROBIOLOGICAL SCI.

XX
PI Nakashima N, Kanamori Y;
DR WPI; 2002-627482/67.

XX
PT Translational activity-promoting higher-order structure of CrPV-like
PT viruses for protein translation when suitably initiated, useful in
PT synthesis of proteins and polypeptides of foreign species for application
PT in drugs.

XX
PS Claim 1; Fig 1-2; 38pp; Japanese.

XX
CC The invention comprises seven RNA sequences (CrPV-like viruses) which
CC have a higher-order structure that sustains translational activity-
CC promoting function. The RNA sequences of the invention are useful in the
CC synthesis of proteins and polypeptides for application in developing and
CC producing drugs. The RNA sequences of the invention are also useful in
CC basic research of protein synthesis and structural analysis by the gene
CC recombinant technique. The present nucleotide represents a Drosophila C
CC virus RNA sequence of the invention

XX
SQ Sequence 201 BP; 61 A; 33 C; 35 G; 0 T; 72 U; 0 Other;

Query Match 65.7%; Score 123.6; DB 6; Length 201;
Best Local Similarity 79.0%; Pred. No. 3.1e-27;

Matches 147; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 3 AAAAUGAUCUUCUCUGUAUAAUACAUAUUUGAGAGUUAUAAUUAACAAGUGUGCUA 62
DB 4 AAGUGAUGAUCUUCUCUGUAUAAUACAUAUUUGAGAGUUAUAAUUAACAAGUGUGCUA 63

QY 63 UUUUUGAUAUUAGGUAUAGCUUAUAGCUUACGUCUCCAGAUCCGCUAGUGCCAGCCCCAC 122
DB 64 UCUUAUAUAUUAGGUAUAGCUUAUAGCUUACGUCUCCAGAUCCGCUAGUGCCAGCCCCAU 123

QY 123 AAUUAUCCAGAGACCCUCUCUGCGUUUUUUCAGAUUAGUGUGGAAAAACCUAAGAAU 182
DB 124 AAUUAUCCAGAGACCCUCUCUGCGUUUAUAGUUAUAGUGUGUUAUAGAAUAAAGAAA 183

QY 183 UUAACCU 188
DB 184 UUAACCU 189

RESULT 5
AB280712
ID AB280712 standard; RNA; 190 BP.

AC AB280712;
XX
DT 15-OCT-2003 (first entry)
DE Black queen-cell virus derived pseudoknot sequence.

XX
DE Black queen-cell virus derived pseudoknot sequence.

XX
KW Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;
KW albumen; impurity; higher-order structure; intergenic region; IGR-IRBS;
KW internal ribosome entry site; ss.

XX
OS Black queen-cell virus.
XX
PN W02003033719-A1.

XX
PD 24-APR-2003.
XX
PF 08-OCT-2002; 2002WO-JP010447.

XX
PR 17-OCT-2001; 2001JP-00319923.
XX
PA (NAG-) NAT INST AGROBIOLOGICAL SCI.

XX
PA (WAK-) WAKENYAKU KK.
XX
PI Nakashima N, Shibuya N, Nishikawa S;
DR WPI; 2003-403230/38.

XX
PT Cell-free protein synthesis means in wheatgerm system to establish
PT overexpression of target gene with base sequence sustaining translation
PT activity and function promotion, for producing useful proteins.

XX
PS Claim 1; Page 34; 39pp; Japanese.

XX
CC The invention relates to to a cell-free protein synthesis system derived
CC from wheatgerm where there is substantial exclusion of wheatgerm embryo
CC albumen impurities. The novel system uses a sequence having a higher-
CC order RNA structure that promotes translation activity. The higher-order
CC sequence is preferably a "pseudoknot", especially derived from a range of
CC viruses (AB280707-AB280713). This sequence represents the "pseudoknot"
CC higher-order sequence from the black queen-cell virus. The sequence is
CC used in a construct which may also include an intergenic region and
CC internal ribosome entry site (IGR-IRBS). The method is applicable in
CC producing useful proteins

XX
SQ Sequence 190 BP; 58 A; 36 C; 40 G; 0 T; 56 U; 0 Other;

Query Match 33.7%; Score 63.4; DB 8; Length 190;
Best Local Similarity 65.9%; Pred. No. 4.2e-03;
Matches 122; Conservative 0; Mismatches 61; Indels 2; Gaps 2;

[illegible]

D _b	2	C A A C A A U G U G A U C U U G C U G U G C G A G G C A A A A U U U G C A C A G A U A A A A U C U G C A A G U A G U G	61
Q _y	60	C U A U U U U U G A U U A G S U A G U A G U A U U A G C U U A G U T C A G A A U G C C U A G - U G G C A G C C	118
D _b	62	C U A U U G U G A A U C A C C G U A C C U A U U D U A G U U U A G C U C C A A G A U C G U G U G A U G C A G C C	121
Q _y	119.	C C A C A A U A U C C A G A A G C C U C U C U G C G A G U U U U U C A G A U U U A G U A G U C G A A A A C U A A G	178
D _b	122	C U A U C A A U A U C U A G A G A A C U G U G C U A G U U U A G A A G A U U A G U A G U C U C U A A C A G A A C	181
Q _y	179	A A A U U	183
D _b	182	A A U U U	186

```

RESULT 7
ABZ80708
ID ABZ80708 standard; RNA; 187 BP.
XX
XX ABZ80708;
AC
XX
XX 15-OCT-2003 (first entry)
DT
XX
XX Himetobi P virus derived pseudoknot sequence.
DE
XX
XX Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;
KW albumen; impurity; higher-order structure; intergenic region; IGR-IRRS;
XX internal ribosome entry site; ss.
XX
XX Himetobi P virus.
OS
XX
XX W02003033719-A1.
PN
XX
XX 24-APR-2003.
PD
XX
XX 08-OCT-2002; 2002WO-JP010447.
PF
XX
XX 17-OCT-2001; 2001JP-00319923.
PR
XX
XX (NAG-) NAT INST AGROBIOLOGICAL SCT.
PA (WAKE-) WAKENYAKU KK.
XX
XX Nakashima N, Shibuya N, Nishikawa S;
PI
XX
XX WPI; 2003-403230/38.
DR
XX
XX Cell-free protein synthesis means in wheatgerm system to establish
PT overexpression of target gene with base sequence sustaining translation
XX activity and function promotion, for producing useful proteins.
XX
XX Claim 1; Page 32; 39pp; Japanese.
PS
XX
XX The invention relates to a cell-free protein synthesis system derived
CC from wheatgerm where there is substantial exclusion of wheatgerm embryo
CC albumen impurities. The novel system uses a sequence having a higher-
CC order RNA structure that promotes translation activity. The higher-order
CC sequence is preferably a "pseudoknot", especially derived from a range of
CC viruses (ABZ80707-ABZ80713). This sequence represents the "pseudoknot"
CC higher-order sequence from the himetobi P virus. The sequence is used in
CC a construct which may also include an intergenic region and internal
CC ribosome entry site (IGR-IRRS). The method is applicable in producing
CC useful proteins
SQ
XX
XX Sequence 187 BP; 53 A; 32 C; 39 G; 0 T; 63 U; 0 Other;
XX
XX
XX Query Match 24.7%; Score 46.4; DB 8; Length 187;
XX Best local Similarity 58.1%; Pred. No. 0.00056;
XX Matches 100; Conservative 0; Mismatches 71; Indels 1; Gaps 1
XX
XX 15 UCCUGUAAATCAAUUUUGAGAGCUAAUAAUUAACAAGUAGUCUUAUUUUUGUAUUUA 74
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 15 UGAUUAAGAAAGUAAAGAAAUUUCUUAUUUAUUUUUUUUUAUAUCUGCUACAUUUUUUAAGAC 74

```


[illegible]

Matches	58;	Conservative	38;	Mismatches	74;	Indels	3;	Gaps	1;
Qy	9	UGAUCUCGUCGUAAUACAUUUUGAGAGGUAAUUAACAAGUAGUCUAUUUUG	68						
Db	7494	TCATCTGGGGCTTAAATATAATTCAATGTAAGAACTGAAGATCATTTTAA	7553						
Qy	69	UAUUAGUUNAGUUAUUAGUUUACUUCACAGAUUCCUAGUUGCAAGCCCAAAUUC	128						
Db	7554	AAATTAAATGTCACATTTTAACTTTTGTGTGAATTTTCTACAGTTCTCAAAATTTT	7613						
Qy	129	CAGAGAACCCUUCUCUGGUGUUCAGAUUAGUAGUCGAAACCUAAGAA	181						
Db	7614	CAGTAAG---TTTCTGCCCTTGTGCAATTATCTGGCAGGCAAAAAAAAAAAAA	7663						
RESULT 12									
ID	AAK7217	standard; DNA; 45017 BP.							
XX	AAK7217;								
AC	AAK7217;								
DT	07-NOV-2001	(first entry)							
XX	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32029.								
XX	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;								
XX	cytostatic; gene therapy; vaccine; metastasis; de.								
OS	Homo sapiens.								
XX	WO200157182-A2.								
XX	09-AUG-2001.								
PD									
PF	17-JAN-2001; 2001WO-US00154.								
XX	31-JAN-2000; 2000US-0179065P.								
PR	04-FEB-2000; 2000US-0180628P.								
PR	24-FEB-2000; 2000US-0184664P.								
PR	02-MAR-2000; 2000US-0186350P.								
PR	16-MAR-2000; 2000US-0189874P.								
PR	17-MAR-2000; 2000US-0190076P.								
PR	18-APR-2000; 2000US-0198123P.								
PR	19-MAY-2000; 2000US-0205515P.								
PR	07-JUN-2000; 2000US-0209467P.								
PR	28-JUN-2000; 2000US-0214886P.								
PR	30-JUN-2000; 2000US-0215135P.								
PR	07-JUL-2000; 2000US-0216647P.								
PR	07-JUL-2000; 2000US-0216880P.								
PR	11-JUL-2000; 2000US-0217487P.								
PR	11-JUL-2000; 2000US-0217496P.								
PR	14-JUL-2000; 2000US-0218290P.								
PR	26-JUL-2000; 2000US-0220963P.								
PR	26-JUL-2000; 2000US-0220963P.								
PR	14-AUG-2000; 2000US-0224518P.								
PR	14-AUG-2000; 2000US-0224519P.								
PR	14-AUG-2000; 2000US-0225213P.								
PR	14-AUG-2000; 2000US-0225214P.								
PR	14-AUG-2000; 2000US-0225267P.								
PR	14-AUG-2000; 2000US-0225267P.								
PR	14-AUG-2000; 2000US-0225268P.								
PR	14-AUG-2000; 2000US-0225270P.								
PR	14-AUG-2000; 2000US-0225447P.								
PR	14-AUG-2000; 2000US-0225757P.								
PR	14-AUG-2000; 2000US-0225758P.								
PR	14-AUG-2000; 2000US-0225759P.								
PR	18-AUG-2000; 2000US-0226219P.								
PR	22-AUG-2000; 2000US-0226681P.								
PR	22-AUG-2000; 2000US-0226686P.								
PR	22-AUG-2000; 2000US-0227182P.								
PR	23-AUG-2000; 2000US-0227009P.								
PR	30-AUG-2000; 2000US-0228924P.								
PR	01-SEP-2000; 2000US-0229287P.								
PR	01-SEP-2000; 2000US-0229343P.								
PR	01-SEP-2000; 2000US-0229344P.								
PR	05-SEP-2000; 2000US-0229509P.								
PR	05-SEP-2000; 2000US-0229513P.								
PR	06-SEP-2000; 2000US-0230437P.								
PR	06-SEP-2000; 2000US-0230438P.								
PR	08-SEP-2000; 2000US-0231242P.								
PR	08-SEP-2000; 2000US-0231243P.								
PR	08-SEP-2000; 2000US-0231244P.								
PR	08-SEP-2000; 2000US-0231413P.								
PR	08-SEP-2000; 2000US-0231414P.								
PR	08-SEP-2000; 2000US-0232080P.								
PR	08-SEP-2000; 2000US-0232081P.								
PR	12-SEP-2000; 2000US-0231968P.								
PR	14-SEP-2000; 2000US-0232397P.								
PR	14-SEP-2000; 2000US-0232398P.								
PR	14-SEP-2000; 2000US-0232399P.								
PR	14-SEP-2000; 2000US-0232400P.								
PR	14-SEP-2000; 2000US-0232401P.								
PR	14-SEP-2000; 2000US-0233063P.								
PR	14-SEP-2000; 2000US-0233064P.								
PR	14-SEP-2000; 2000US-0233065P.								
PR	21-SEP-2000; 2000US-0234223P.								
PR	21-SEP-2000; 2000US-0234274P.								
PR	25-SEP-2000; 2000US-0234597P.								
PR	25-SEP-2000; 2000US-0234598P.								
PR	25-SEP-2000; 2000US-0234599P.								
PR	26-SEP-2000; 2000US-0235484P.								
PR	27-SEP-2000; 2000US-0235834P.								
PR	27-SEP-2000; 2000US-0235836P.								
PR	29-SEP-2000; 2000US-0236327P.								
PR	29-SEP-2000; 2000US-0236367P.								
PR	29-SEP-2000; 2000US-0236368P.								
PR	29-SEP-2000; 2000US-0236369P.								
PR	29-SEP-2000; 2000US-0236370P.								
PR	02-OCT-2000; 2000US-0236802P.								
PR	02-OCT-2000; 2000US-0237037P.								
PR	02-OCT-2000; 2000US-0237038P.								
PR	02-OCT-2000; 2000US-0237039P.								
PR	02-OCT-2000; 2000US-0237040P.								
PR	13-OCT-2000; 2000US-0239935P.								
PR	13-OCT-2000; 2000US-0239937P.								
PR	20-OCT-2000; 2000US-0240960P.								
PR	20-OCT-2000; 2000US-0241221P.								
PR	20-OCT-2000; 2000US-0241785P.								
PR	20-OCT-2000; 2000US-0241786P.								
PR	20-OCT-2000; 2000US-0241787P.								
PR	20-OCT-2000; 2000US-0241808P.								
PR	20-OCT-2000; 2000US-0241809P.								
PR	20-OCT-2000; 2000US-0241826P.								
PR	01-NOV-2000; 2000US-0244617P.								
PR	08-NOV-2000; 2000US-0246474P.								
PR	08-NOV-2000; 2000US-0246475P.								
PR	08-NOV-2000; 2000US-0246476P.								
PR	08-NOV-2000; 2000US-0246477P.								
PR	08-NOV-2000; 2000US-0246478P.								
PR	08-NOV-2000; 2000US-0246532P.								
PR	08-NOV-2000; 2000US-0246533P.								
PR	08-NOV-2000; 2000US-0246609P.								
PR	08-NOV-2000; 2000US-0246610P.								
PR	08-NOV-2000; 2000US-0246611P.								
PR	08-NOV-2000; 2000US-0246613P.								
PR	17-NOV-2000; 2000US-0249207P.								
PR	17-NOV-2000; 2000US-0249208P.								
PR	17-NOV-2000; 2000US-0249209P.								
PR	17-NOV-2000; 2000US-0249210P.								
PR	17-NOV-2000; 2000US-0249211P.								
PR	17-NOV-2000; 2000US-0249212P.								

PR	01-SEP-2000	2000US-022934.33
PR	01-SEP-2000	2000US-022934.46
PR	01-SEP-2000	2000US-022934.55
PR	05-SEP-2000	2000US-022950.99
PR	05-SEP-2000	2000US-022951.33
PR	06-SEP-2000	2000US-023043.78
PR	06-SEP-2000	2000US-023043.88
PR	08-SEP-2000	2000US-023124.92
PR	08-SEP-2000	2000US-023124.93
PR	08-SEP-2000	2000US-023124.94
PR	08-SEP-2000	2000US-023141.33
PR	08-SEP-2000	2000US-023208.60
PR	08-SEP-2000	2000US-023208.61
PR	12-SEP-2000	2000US-023196.68
PR	14-SEP-2000	2000US-023223.77
PR	14-SEP-2000	2000US-023233.88
PR	14-SEP-2000	2000US-023233.99
PR	14-SEP-2000	2000US-023242.47
PR	14-SEP-2000	2000US-023493.77
PR	14-SEP-2000	2000US-023240.13
PR	14-SEP-2000	2000US-023306.33
PR	14-SEP-2000	2000US-023306.44
PR	14-SEP-2000	2000US-023306.55
PR	21-SEP-2000	2000US-023422.33
PR	21-SEP-2000	2000US-023422.34
PR	21-SEP-2000	2000US-023422.35
PR	25-SEP-2000	2000US-023459.77
PR	25-SEP-2000	2000US-023459.78
PR	25-SEP-2000	2000US-023459.79
PR	25-SEP-2000	2000US-023498.84
PR	27-SEP-2000	2000US-023589.44
PR	27-SEP-2000	2000US-023589.45
PR	27-SEP-2000	2000US-023589.46
PR	29-SEP-2000	2000US-023633.77
PR	29-SEP-2000	2000US-023633.78
PR	29-SEP-2000	2000US-023633.79
PR	29-SEP-2000	2000US-023636.68
PR	29-SEP-2000	2000US-023636.69
PR	29-SEP-2000	2000US-023637.00
PR	02-OCT-2000	2000US-023690.22
PR	02-OCT-2000	2000US-023703.77
PR	02-OCT-2000	2000US-023703.78
PR	02-OCT-2000	2000US-023703.79
PR	02-OCT-2000	2000US-023704.00
PR	13-OCT-2000	2000US-023939.53
PR	13-OCT-2000	2000US-023939.54
PR	20-OCT-2000	2000US-024096.00
PR	20-OCT-2000	2000US-024122.13
PR	20-OCT-2000	2000US-024122.14
PR	20-OCT-2000	2000US-024178.53
PR	20-OCT-2000	2000US-024178.54
PR	20-OCT-2000	2000US-024178.55
PR	20-OCT-2000	2000US-024180.88
PR	20-OCT-2000	2000US-024180.89
PR	01-NOV-2000	2000US-024461.77
PR	01-NOV-2000	2000US-024647.44
PR	08-NOV-2000	2000US-024647.53
PR	08-NOV-2000	2000US-024647.54
PR	08-NOV-2000	2000US-024647.55
PR	08-NOV-2000	2000US-024652.77
PR	08-NOV-2000	2000US-024652.78
PR	08-NOV-2000	2000US-024652.79
PR	08-NOV-2000	2000US-024653.22
PR	08-NOV-2000	2000US-024653.23
PR	08-NOV-2000	2000US-024653.24
PR	08-NOV-2000	2000US-024661.10
PR	08-NOV-2000	2000US-024661.11
PR	11-NOV-2000	2000US-024661.33
PR	11-NOV-2000	2000US-024620.77
PR	11-NOV-2000	2000US-024920.88
PR	11-NOV-2000	2000US-024920.89
PR	11-NOV-2000	2000US-024921.00
PR	11-NOV-2000	2000US-024921.21
PR	11-NOV-2000	2000US-024921.22

QY		63	UUUUGUUUUUUAGUAGUAUCUAAUUUAGCUUACGUCACAGAGUCGUAGUGGACGCCGCAC	122
Db		150	TTTTAATATAATTAGGTAACTATTAGTTTACTGTTTCAGAGTGCCATTGGCAGGCCCAT	209
QY		123	AUAUUCGAGAAGCCUCUCUGCGGCUUUUUUCAUAUUGAGUAGUCGAAAAACUAAAGAAU	182
Db		210	AATATCCAGGACAACCCTCTGCTCTTATATGAATTAGTGTGCAITTTGAAATAGAAA	269
QY		183	UUACCU 188	
Db		270	TAACTT 275	
RESULT 2				
LOCUS		COI53454	486 bp mRNA linear EST 05-OCT-2004	
DEFINITION		EN2010.5prime Exelixis FlyTag MN08 Bluescript Drosophila melanogaster cDNA clone EN2010 5, mRNA sequence.		
ACCESSION		COI53454		
VERSION		COI53454.1	GI:48907455	
SOURCE		EST.		
ORGANISM		Drosophila melanogaster (fruit fly)		
REFERENCE		Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 486) Nakanishi,M., Zhang,C., Peterson,E., Laufer,A., Leung,W., Platt,D. and Swimmer,C. Exelixis FlyTag EST Project MN08 Library Unpublished (2004) Contact: Stapleton, M. BDGP Lawrence Berkeley National Lab One Cyclotron Rd, Berkeley, CA 94720, USA Fax: 510 486 6798 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu Plate: EN.20 row: A column: 10 High quality sequence stop: 412. Location/Qualifiers 1..486 /organism="Drosophila melanogaster" /mol_type="mRNA" /db_xref="taxon:7227" /clone="EN2010" /cell_line="mbn2" /clove_lib="Exelixis FlyTag MN08 Bluescript" /note="Vector: Bluescript; Site_1: NotI; Site_2: XhoI; oligodt primed from LPS induced mbn2 cell line."		
FEATURES				
source				
ORIGIN				
Query Match		66.6%; Score 125.2; DB 7; Length 486;		
Best local Similarity		51.1%; Pred. 1.6e-24;		
Matches		95; Conservative 53; Mismatches 38; Indels 0; Gaps 0;		
QY		3	AAAAGUAGUUCUGUCUGUAAAUAACAUAUUUGAAGAGUUAUAAAUCAAAGUAGUCUA	62
Db		68	AAAGATGATCTTGCTCTTATACAATTTTGAAGAGTTATTAAGAAGAGTAGTGCTA	127
QY		63	UUUUUGAUUUUAGUAGCUAAUUUAGUUAAGUUCGAGUAGUCUGGACGCCCAC	122
Db		128	TTTTTAATATAATTAGGTAACTATTAGTTTACTGTTTCAGAGTGCCATTGGCAGGCCCAT	187
QY		123	AUAUUCGAGAAGCCUCUCUGCGGCUUUUUUCAUAUUGAGUAGUCGAAAAACUAAAGAAU	182
Db		188	AATATCCAGGACAACCCTCTGCTCTTATATGAATTAGTGTGCAITTTGAAATAGAAA	247
QY		183	UUACCU 188	
Db		248	TAACTT 253	

RESULT 3	
LOCUS	CO153765
DEFINITION	CO153765 490 bp mRNA linear EST 05-OCT-2004
ACCESSION	EN02756.Sprine Exelixis FlyTag MN08 Bluescript Drosophila melanogaster cDNA clone EN02756 5, mRNA sequence.
VERSION	CO153765
KEYWORDS	CO153765.1 GI:48907766
SOURCE	EST.
ORGANISM	Drosophila melanogaster (fruit fly)
REFERENCE	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 490)
AUTHORS	Nakanishi,M., Muzong,C., Peterson,E., Laufer,A., Leung,W., Platt,D. and Swimmer,C.
TITLE	Exelixis FlyTag EST Project MN08 Library
JOURNAL	Unpublished (2004)
COMMENT	Contact: Stapleton, M. BDGP Lawrence Berkeley National Lab One Cyclotron Rd, Berkeley, CA 94720, USA Fax: 510 486 6798 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu Plate: EN.27 row: E column: 8 High quality sequence stop: 401. Location/Qualifiers 1..490 /organism="Drosophila melanogaster" /mol_type="mRNA" /db_xref="taxon:7227" /clone="EN02756" /cell_line="mbn2" /clone_lib="Exelixis FlyTag MN08 Bluescript" /note=Vector: pBluescript; Site 1: NotI; Site 2: XhoI; oligodt primed from LPS induced mbn2 cell line."
ORIGIN	
Query Match	66.6%; Score 125.2; DB 7; Length 490;
Best Local Similarity	51.1%; Pred. No. 1.6e-24;
Matches	95; Conservative 53; Mismatches 38; Indels 0; Gaps 0;
QY	3 AAAAUGAGCUUUGGUAUUAAUUUGAGGCGUUAUUAAUUACAAGUCUCUA 62
Db	104 AAGGTGGATCTTGCTTCCTTAATCAATTGGAGAGTTATTAAGAAGATGTGCTA 163
QY	63 UUUUUGAUUUAGGUAGCUAUUUAGCUUUACGUCCAGAGUCCGUAUGGCACGCCAC 122
Db	164 TTTTATAATTAAGTAGTAACTATTGAATTACGTTTCAGAGTGCCTATTTGGCAGCCCCAT 223
QY	123 AAANAUCAGGAACCCUUCUUCUGGCUUUUUUCAUAUNAGUAGGAAAAACCUAAGAAU 182
Db	224 AATATCCAGGACACCCTCTGCTTCTTAATATGATTAGGTTCATTGAATAAGAAA 283
QY	183 UVACCU 188
Db	284 TAACCT 289
RESULT 4	
LOCUS	CO337794
DEFINITION	CO337794 539 bp mRNA linear EST 05-OCT-2004
ACCESSION	EN15317.Sprine Exelixis FlyTag MN08 Bluescript Drosophila melanogaster cDNA clone EN15317 5, mRNA sequence.
VERSION	CO337794
KEYWORDS	CO337794.1 GI:49398069
SOURCE	EST.
ORGANISM	Drosophila melanogaster (fruit fly) Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 539)

Db 426 TATTTAGCTTAACTATTATTAATTAGCTGTTTATGAACCGCAGAGG 476

RESULT 9 455 bp mRNA linear EST 15-OCT-2001
 BB746609
 BB746609
 DEFINITION BB746609 RIKEN full-length enriched, adult male kidney Mus musculus
 cDNA clone F530101N20 3', mRNA sequence.
 ACCESSION BB746609
 VERSION BB746609.1 GI:16149546
 KEYWORDS EST.
 ORGANISM Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.
 1 (bases 1 to 455)
 REFERENCE 1 Akimura,T., Arikawa,T., Carninci,P., Furuno,M., Hasegaki,T.,
 Hayatsu,N., Hiramoto,K., Hiroaka,T., Hirozane,T., Imotani,K.,
 Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Komoto,H., Kouda,M.,
 Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
 Saeki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaki-Akahira,S.,
 Tanaka,T., Tomaru,A., Toyo,T., Watahiki,A., Yasunishi,A.,
 Muramatsu,M. and Hayashizaki,Y.
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
 2001)
 JOURNAL Unpublished (2001)
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Saito-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Komoto,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Itawa,M., Ohara,E.,
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
 Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
 and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Komoto,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
 Sugahara,Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-288 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 e mouse tissues.
 Location/Qualifiers
 1. 455
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="F530101N20"
 /sex="male"
 /tissue_type="kidney"
 /dev_stage="adult"
 /lab_host="SOLR"
 /clone_id="RIKEN full-length enriched, adult male kidney"
 /note="Site 1: XhoI; Site 2: SacI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken

[illegible]

[illegible]

ORGANISM	Brassica oleracea
SOURCE	Brassica oleracea
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons; rosoids; eurosid II; Brassicales; Brassicaceae; Brassica.
AUTHORS	(bases 1 to 777) Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S., Utecherack,T.R., Mortman,J.R., White,O.R. and Town,C.D.
TITLE	Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis
JOURNAL	Genome Res. 15 (4), 487-495 (2005)
PUBMED	15603490
COMMENT	Other_GSSs: BOGUU79TR Contact: Chris Town TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: TP Class: sheared ends.
FEATURES	location/Qualifiers 1..777 /organism="Brassica oleracea" /mol_type="genomic DNA" /strain="TO100DH3" /db_xref="taxon:3712" /cclone="BOGUU79" /clone_1lb="BOGU" /vector="PHOS1; Site 1: BstXI, 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"
ORIGIN	
Query Match	20.1%; Score 37.8; DB 9; Length 777;
Best Local Similarity	33.0%; Pred. No. 7.5;
Matches	32; Conservative 28; Mismatches 37; Indels 0; Gaps 0;
DY	15 UGCUGUAUUAAUUCAGUUUUUGAGGCUUAUAUUACAGUAGUGCUAUUUUUUGUUAUUUA 74 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : Db 305 TTCAATTAAATACTGTCTTGACAGACGGAAGATTCAAACTCGATTAATTTTGATCTC 364
Dy	75 GUUAGCUAUUUAGUCUUUACGUCCAGAGGCCUAGU 111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : Db 365 GTTAGCTCTTAGCTTACACTGTACAGAATCTGT 401
RESULT 13	
LOCUS	CO338064 453 bp mRNA linear EST 05-OCT-2004
DEFINITION	EN15862.5prime Exelixis fLYTag MN08 Bluescript Drosophila melanogaster cDNA clone EN15862.5, mRNA sequence.
ACCESSION	CO338064
VERSION	CO338064.1
KEYWORDS	GI:49398359
SOURCE	EST.
ORGANISM	Drosophila melanogaster (fruit fly)
PROSPEC	Drosophila melanogaster
EUKARYOTA	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 453) Nakanishi,M., Muzong,C., Peterson,E., Lauffer,A., Leung,W., Platt,D. and Swimmer,C.
AUTHORS	Exelixis fLYTag EST Project MN08 Library
TITLE	Unpublished (2004)
JOURNAL	Contact: Stapleton, M.
COMMENT	BDBG Lawrence Berkeley National Lab One Cyclotron Rd, Berkeley, CA 94720, USA Fax: 510 486 6798 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu Plate: EN.158 row: F column: 2 High quality sequence stop: 399.


```

FEATURES
  source
    location/Qualifiers
      1..453
      /organism="Drosophila melanogaster"
      /mol_type="mRNA"
      /db_xref="taxon:7227"
      /clone="EN15862"
      /cell_line="mbn2"
      /clone_lib="Exelixis FlyTag MN08 Bluescript"
      /note="Vector: pBluescript; Site 1: NotI; Site 2: XhoI;
      oligodr primed from LPS induced mbn2 cell line."
  ORIGIN
    Query Match      20.0%; Score 37.6; DB 7; Length 453;
    Best Local Similarity 50.0%; Pred. No. 7.8;
    Matches 34; Conservative 15; Mismatches 19; Indels 0; Gaps 0;
  QY
    121 ACAAUUACGAGGAGCCGUCGCGGUGUUUUCAGAUUAGGUGAGGAAAAACCUAGAA 180
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
  Db 1 ATAATATCGAGACACCTCTCTGCTTCTTATATGATTAAGTTGTCATTGAAATAGAA 60
      |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
  QY 181 AUUUAUCCU 188
      |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
  Db 61 AATAACCT 68
      |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
  RESULT 14
  C2696233 452 bp DNA linear GSS 25-JUL-2005
  LOCUS OC_Ba0009L24.f OC_Ba Oryza coarctata genomic clone OC_Ba0009L24
  DEFINITION 5'__genomic survey sequence.
  ACCESSION C2696233
  VERSION C2696233.1 GI:71095381
  KEYWORDS GSS.
  ORGANISM Oryza coarctata (Porteresia coarctata)
  Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzaceae; Oryza.
  1 (bases 1 to 452)
  Klu,H., Collura,K., Misosaki,M., Byrne,M., Stum,D., Smart,D.,
  Rao,K., Luo,M., Jettly,R., Kudrna,D., Muller,C., Soderlund,C. and
  Wing,R.
  OMP (Oryza Map Alignment Project) - Arizona Genomics Institute
  Unpublished (2005)
  Contact: Rod A. Wing
  Arizona Genomics Institute
  University of Arizona
  Forbes Building Room 303, Tucson, AZ 85721-0036, USA
  Tel: 520 626 9595
  Fax: 520 621 1259
  Email: rwing@genome.arizona.edu
  PCR primers
  FORWARD: TAA TAC GAC TCA CTA TAG GG
  BACKWARD: CAC TCA TTA GGC ACC CCA
  Plate: 0009 row: L column: 24
  Seq primer: TAA TAC GAC TCA CTA TAG GG
  Class: BAC ends.
  location/Qualifiers
    1..452
    /organism="Oryza coarctata"
    /mol_type="genomic DNA"
    /db_xref="taxon:77588"
    /clone="OC_Ba0009L24"
    /tissue_type="leaves"
    /dev_stage="mature"
    /lab_host="DH10B"
    /clone_lib="OC_Ba"
    /note="Vector: pGIRIBAC1, Site_1: HindIII, Site_2: HindIII"
  FEATURES
  source
    19.7%; Score 37; DB 10; Length 452;
    Best Local Similarity 36.6%; Pred. No. 12;
    Matches 37; Conservative 24; Mismatches 40; Indels 0; Gaps 0;
  ORIGIN
    Query Match      19.7%; Score 37; DB 10; Length 452;
    Best Local Similarity 36.6%; Pred. No. 12;
    Matches 37; Conservative 24; Mismatches 40; Indels 0; Gaps 0;
  ORIGIN

```

Oy		4	AAANGUAGUCUGGUCGUAAAUAACAUUUUGAGAGGUTAAUAUAUUUACAAGUAGUCGAU	63	
Db		123	AGATTTCCTCAACTTGTGTATTAATAACAGAGCGTGAAAAAACAAGATATAGTGAT	64	
Oy		64	UUUGUUAUUUAGGUGUAGCUAUUUUAGCCUUAUGCUCAGAGAU	104	
Db		63	TTTGTAGTAGTGGTTAGGTTGGTGTAGTGGTCCAGAT	23	
RESULT_15	AJ702451	485 bp	mRNA	linear	EST 28-JAN-2005
LOCUS	AJ702451				
DEFINITION	AJ702451 ZAAE floral spathe MP1Z-ADIS-062 cDNA Zantedeschia				
ACCESSION	Aethiopica cDNA clone J-18-4, mRNA sequence.				
VERSION	AJ702451				
KEYWORDS	AJ702451.1 GI:58318066				
SOURCE	EST.				
ORGANISM	Zantedeschia aethiopica Zantedeschia aethiopica Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Araceae; Aroidae; Zantedeschieae; Zantedeschia. 1 (bases 1 to 485) Piquet,M.C., Weider,M., Lino-Neto,T., Barbeta,C., Tavares,R.M., Weissenhaar,B., Pais,M.S. and Palme,K. Gene expression analysis during Zantedeschia aethiopica floral spathe senescence and senescence-inhibition using EST sequencing and cDNA arrays Unpublished (2005) Contact : Piquet MC Institut fuer Biologie II Albert-Ludwigs-Universitaet Freiburg Schauzlestr. 1, D-79104 Freiburg, GERMANY. Location/Qualifiers 1..485 /organism="Zantedeschia aethiopica" /mol_type="mRNA" /db_xref="taxon:69721" /clone="J-18-4" /tissue_type="floral spathe" /dev_stage="floral spathe whitening and regreening" /lab_host="Escherichia coli strain TOP10" /clone_id="ZAAE floral spathe MP1Z-ADIS-062 cDNA" /note="Vector: pSPORT1; cDNA synthesis was performed using SuperScript system (Invitrogen) and directionally-cloned (Sal I at the 5'-end and Not I at the 3'-end) into pSPORT1 (UI2390.1) country: Portugal:Braga"				
ORIGIN					
Query Match	19.7%; Score 37; DB 1; Length 485;				
Best Local Similarity	30.1%; Pred. No. 12;				
Matches	28; Conservative 30; Mismatches 35; Indels 0; Gaps 0;				
Oy		6	AUGGAGCUCUGGUCGUAAAUAACAUUUUGAGAGGUTAAUAUAUUUACAAGUAGUCGAU	65	
Db		228	ATAATATTGCCTCGCAAATGCATATTAGTATGAACATCATTTGATAGGAATTTATTT	287	
Oy		66	UUUGUUAUUUAGGUGUAGCUAUUUUAGCCUUAUGCUC	98	
Db		288	GTTGATATACATTTTCTTAGATAGATGATTTATTTTC	320	

Search completed: December 22, 2005, 15:21:02

Job time : 1971.58 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 07:55:17 ; Search time 79.0668 Seconds
(without alignments)
4226.575 Million cell updates/sec

Title: US-10-088-750C-4

Perfect score: 188
Sequence: 1 caaaagaugaucugcug.....aaacuaagaauuaccu 188

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/1/COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5/COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A/COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B/COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H/COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS/COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP/COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE/COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	125.2	66.6	1345	US-09-270-767-1072
2	125.2	66.6	1345	US-09-270-767-1072
3	38.6	20.5	54444	US-09-949-016-17344
4	34.2	18.2	601	US-09-949-016-56332
5	34.2	18.2	173992	US-09-949-016-13379
6	33.2	17.7	108169	US-09-949-016-12898
7	33.2	17.7	108169	US-09-949-016-15907
8	32.6	17.3	25769	US-09-949-016-17154
9	32.6	17.3	67876	US-09-949-016-12332
10	32.4	17.2	601	US-09-949-016-192590
11	32.4	17.2	1458	US-09-081-686-3
12	32.4	17.2	2279	US-09-543-681A-1309
13	32.2	17.1	1407	US-09-543-681A-1449
14	32.2	17.1	1407	US-09-543-681A-1449
15	32.2	17.1	115963	US-09-949-016-12332
16	32.2	17.0	601	US-09-949-016-183399
17	32.2	17.0	140224	US-09-949-016-17002
18	31.8	16.9	601	US-09-949-002-4496
19	31.8	16.9	185765	US-09-949-002-674
20	31.8	16.9	185766	US-09-949-002-707
21	31.8	16.9	185959	US-09-949-016-13125
22	31.6	16.8	436	US-09-513-999C-18636
23	31.6	16.8	601	US-09-949-016-170092
24	31.6	16.8	10330	US-10-001-189-66

25	31.6	16.8	35609	US-09-949-016-17370	Sequence 17370, A
26	31.6	16.8	96327	US-09-949-016-16541	Sequence 16541, A
27	31.4	16.7	3484	US-09-308-090-1	Sequence 1, Appl1
28	31.4	16.7	3484	US-09-380-090A-1	Sequence 1, Appl1
29	31.4	16.7	10330	US-10-001-189-68	Sequence 68, Appl1
30	31.2	16.6	601	US-09-949-016-59922	Sequence 59922, A
31	31.2	16.6	601	US-09-949-016-59923	Sequence 59923, A
32	31.2	16.6	601	US-09-949-016-86428	Sequence 86428, A
33	31.2	16.6	52655	US-09-949-016-13482	Sequence 13482, A
34	31.2	16.6	422592	US-09-949-016-12053	Sequence 12053, A
35	31	16.5	87352	US-09-949-016-12721	Sequence 12721, A
36	31	16.5	87352	US-09-949-016-15692	Sequence 15692, A
37	31	16.5	87352	US-09-949-016-15692	Sequence 15693, A
38	31	16.5	87352	US-09-949-016-15693	Sequence 1, Appl1
39	31	16.5	1830121	US-09-557-884-1	Sequence 1, Appl1
40	31	16.5	1830121	US-09-543-990A-1	Sequence 1, Appl1
41	31	16.5	1830121	US-10-158-865-1	Sequence 12, Appl1
42	30.8	16.4	441	US-08-568-310D-12	Sequence 12, Appl1
43	30.8	16.4	441	US-09-270-455-12	Sequence 168, App
44	30.8	16.4	1332	US-09-769-787-168	Sequence 213, App
45	30.8	16.4	1332	US-09-769-787-213	

ALIGNMENTS

```

RESULT 1
US-09-270-767-1072
; Sequence 1072, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1072
; LENGTH: 1345
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1072

Query Match 66.6%, Score 125.2, DB 3; Length 1345;
Best Local Similarity 51.1%, Pred. No. 2.3e-30;
Matches 95; Conservative 53; Mismatches 38; Indels 0; Gaps 0;

QY 3 AAAUGAUGUCUGUCUGUAUAUACAUAUUUGAGAGUGUAUAUAUACAAGUGUGUA 62
DB 520 AAGATGATCTTCTCTCTTATACATTTTGAGAGTTATAGAGAGAGAGAGAGAG 579
QY 63 UUUUGAUAUUUGAUGAUGAUAUUUGAUGAUGAUGAUGAUGAUGAUGAUGAUGA 122
DB 580 TTTTAATATATAGGTTAATATATATATATATATATATATATATATATATATAT 639
QY 123 AAUUGCAGAGAGCCGUCUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGA 182
DB 640 AATATCAGAGACCCCTCTCTCTTATATATATATATATATATATATATATAT 699
QY 183 UUAUCCU 188
DB 700 TAACTT 705

RESULT 2
US-09-270-767-16354
; Sequence 16354, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094

```



```

:
: CURRENT APPLICATION NUMBER: US/09/270,767
:
: CURRENT FILING DATE: 1999-03-17
:
: NUMBER OF SEQ ID NOS: 62517
:
: SOFTWARE: Patent In Ver. 2.0
:
: SEQ ID NO 16354
:
: LENGTH: 1345
:
: TYPE: DNA
:
: ORGANISM: Drosophila melanogaster
:
: US-09-270-767-16354

```

Query Match	66.6%;	Score 125.2;	DB 3;	Length 1345;
Best Local Similarity	51.1%;	Pred. No. 2.3e-30;		
Matches 95;	Conservative 53;	Mismatches 38;	Indels 0;	Gaps 0;

Qy 3 AAAAUGGUCUUCGUGAAAUAUCAUUUGAGGUGUAAUUAUUAACAAGUGUGCA 62
|||:::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 520 AAGATGTGATCTTGCTTACCAATTGGAGGGTTATATAAGAAGNAGTGTGCTA 579

63 UUUUGUAAUAGGUGUAGCUUUUAGCUUACGUUCAGAGUGGCUAGUGGACGCCAC 122
 580 TTTTATAATTAGTTACTATTGATTTTACTGTTCAGAGGCTATTGGCAGCCCAT 639

0Y AAATCCAGAAAGCTCCTCCGCGGGUUUUUUCAGAUAGAGUAUGAAAAACCUAAGAAGU 182
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 640 AATATCCAGAACACCCTCTCTGTTCTTATGATTGAATTAGTGTTGTCATTTCGATATAGAAA 699

QY	183	UUACCU	188
	:	:	
Db	700	TAACCT	705

RESULT 3
US-09-949-016-17344/C
Serials 17344 180140044m mc/00000007

; PATENT NO. 6812539
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: VENTER, J. Craig et al.
 ;
 ; TITLE OF INVENTION: POLYNUCLEOTIDES IN VIVO GENE ASSOCIATED

```

1  TITLE OF INVENTION: WITH HUMAN DATABASE, METHODS OF
2  FILE REFERENCE: CLO01307
3  CURRENT APPLICATION NUMBER: US/09/949,016
4  CURRENT FILING DATE: 2000-04-14
5  THERE

```

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSeq for Windows Version 4.0

```

; LENGTH: 54444
;
; TYPE: DNA
;
; ORGANISM: Human

```

Query Match	20.5%	Score 39.6	DB 3	Length 54444
Best Local Similarity	33.5%	Pred. No. 0.079		

9 UGAUCUUCGCTUGAAAUACAUAUUUGAGAGGUAUAUAAUAACAAGUAGUCUAUUUUUG 68

69 UAUUUAGCUUAGCUAUUUAGCUUAGCUUCCAGAGUGCCUAGUGGACAGCCCCACAUAUUC 128

QY 129 CAGGAAGCCCTCUCGCGGUTUUUACAGAUUAGUGAUGAAAAACCUAAGAAA 181
||||| : : ||| : : ||| : : ||| : : |||

RESULT 4
US-09-949-016-56332
; Sequence 56332, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; ATTORNEY FOR INVENTOR: J. Craig Venter

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949, 016
 ; CURRENT FILING DATE: 2009-09-01

```

/ PRIOR APPLICATION NUMBER: 60/441,155
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-02

```

```

1  PRIOR REGISTRATION NUMBER: 007421,750
2  PRIOR FILING DATE: 2000-09-08
3  NUMBER OF SEQ ID NOS: 207012
4  SOFTWARE: PAIR-SEO for Windows Version 4.0

```

/ CNA ID: NO 26532
 /
 / LENGTH: 601
 /
 / TYPE: DNA
 /
 / ORGANISM: Human

Query Match	18.2%	Score 34.2	DB 3	length 601
Best Local Similarity	32.4%	Pred NO	0.34	

3 AAAAUGGUAUCUUGGUAUAAACAUAUUUUGAGAGUUAAUAAUUAACAAGUAUCUA 62

OY

63 TTUTUGAUVU 73

: : : : :

RESULT 5

```

: Sequence 13379, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:

```

FILE REFERENCE: CLO001307

```

;
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
;

```

;
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ;

```
; SOFTWARE: FastSeq IoT Windows Version 4.0  
; SEO ID NO 13379  
; LENGTH: 173992  
mimo
```

US-09-949-016-13379
ORGANISM: Human
Query: March 16 08 09:00:00

Best Local	32.44%	Pred. NO. 3.5;
Matches	23;	Mismatches
Conservative	25;	Indels
		Gaps
		0;

DB 158503 AAAAAGGACACCTCTGGGTGATAAAAATTAGTAACTTAACTCAACACCAATAAATTTT 158544

QY 63 UUUUUGUAUU 73

Db 158543 TTTT TTTT TTTT 158533

RESULT 6
US-09-949-016-12898/c
Sequence 12898, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR FILING DATE: 2000-10-20

PRIOR FILING DATE: 2000-10-03

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 12898

LENGTH: 108169

TYPE: DNA

ORGANISM: Human

US-09-949-016-12898

Query Match 17.7%, Score 33.2, DB 3, Length 108169;
Best Local Similarity 30.6%, Pred. No. 5.7;
Matches 41, Conservative 30, Mismatches 63, Indels 0, Gaps 0;

Qy 19 UGUAAUACAUAUUUGAGAGUUAUAUAUACAAGUAGUCUAUUUUUGUAUUAGCU 78

Db 57934 TGTAAATGGCATTATGTTTAAATCTCAATCTTCTTCTTGTGTAACATGAA 57875

Qy 79 AGCUAUUAGCUUACGUCUCCAGAGUCCUAGUAGUAGUAGUAGUAGUAGUAGU 138

Db 57874 AGCAGTGAATCTTGTGTATTAAGCTTACCTGCACTTCACTATCTACTTATAG 57815

Qy 139 UCUCUGCGUUUUU 152

Db 57814 TTTTCAGAGATTTT 57801

RESULT 7
US-09-949-016-15907/c
Sequence 15907, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR FILING DATE: 2000-10-20

PRIOR FILING DATE: 2000-10-03

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 15907

LENGTH: 108169

TYPE: DNA

ORGANISM: Human

US-09-949-016-15907

Query Match 17.7%, Score 33.2, DB 3, Length 108169;
Best Local Similarity 30.6%, Pred. No. 5.7;
Matches 41, Conservative 30, Mismatches 63, Indels 0, Gaps 0;

Qy 19 UGUAAUACAUAUUUGAGAGUUAUAUAUACAAGUAGUCUAUUUUUGUAUUAGCU 78

Db 57934 TGTAAATGGCATTATGTTTAAATCTCAATCTTCTTGTGTAACATGAA 57875

Qy 79 AGCUAUUAGCUUACGUCUCCAGAGUCCUAGUAGUAGUAGUAGUAGUAGUAGU 138

Db 57874 AGCAGTGAATCTTGTGTATTAAGCTTACCTGCACTTCACTATCTACTTATAG 57815

Qy 139 UCUCUGCGUUUUU 152

Db 57814 TTTTCAGAGATTTT 57801

RESULT 8
US-09-949-016-17154
Sequence 17154, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR FILING DATE: 2000-10-20

PRIOR FILING DATE: 2000-10-03

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 17154

LENGTH: 25769

TYPE: DNA

ORGANISM: Human

US-09-949-016-17154

Query Match 17.3%, Score 32.6, DB 3, Length 25769;
Best Local Similarity 31.6%, Pred. No. 5;
Matches 30, Conservative 26, Mismatches 39, Indels 0, Gaps 0;

Qy 2 AAAAUGGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 61

Db 15626 AAAAATGTAATTAATGTAATTAATGTAATTAATGTAATTAATGTAATTAATG 15685

Qy 62 AUUUUGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 96

Db 15686 GATTAGATTAATTAAGATATCTTTTATTTTACTT 15720

RESULT 9
US-09-949-016-17232/c
Sequence 17232, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR FILING DATE: 2000-10-20

PRIOR FILING DATE: 2000-10-03

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 17232

LENGTH: 67876


```

; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17232

Query Match      17.3%   Score 32.6;  DB 3;  Length 67876;
Best Local Similarity 32.3%   Pred. No. 7.3;
Matches 41;  Conservative 27;  Mismatches 59;  Indels 0;  Gaps 0;

QY      5  A AUGUGAUCUGCGUCUGUAAUACAUUUUGAGAGUUAAUAAUUUACAAGUAGUCUATU 64
      |||  |||  ::  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::
Db      62966 ATTGCTAATATATTATTAACACTTAAAGAAAGTTTACATATTTGAAAAAGGCTATT 62907

QY      65  UUUUGAUUUUAGUGAGCUAUUUUAGCUUUAAGUUCACAGAUUGCCUAGUGACGCCCAAA 124
      ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db      62906 GTCTAATTCACCTTATTATGTCTACATACATGATGACAACTTATTACAAGACAATT 62847

QY      125  UAUCCAG 131
      ::  |||  |||
Db      62846 TACCCAG 62840

```

```

RESULT 10
US-09-949-016-192590/c
; Sequence 192590, Application US/03949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 192590
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; IS-09-949-016-192590

```

	Query Match	17.2%	Score 32.4	DB 3	Length 601
	Best Local Similarity	33.3%	Pred. No.1,3		
	Matches	38	Conservative	25	Mismatches 51
					Indels 0
					Gaps 0
QY	18	UUGUAAUACA	UUUUUGAGAGUUAUAAUAAACAAGUAGCUAUUUUUGAUUUUAGU	77	
DB	592	TTATTA	CTCACTTTAAAAAGATTTCACATATTGGAAAAAGGGCTATTGTGCAATTAATCACT	533	
QY	78	UAGCUAUUUAGCUUACGUUCCAGAGUCCUAGUGGACGCCCCACAUAUCCAG	131		
b	532	TTATTATGTCTACATCATGTGTAGCAACCTATTACAAGAACAATTCTTAGCCAG	479		

US-09-081-686-1
RESULT 11
Sequence 1, Application US/09081686-1
Patent No. 6162619
GENERAL INFORMATION:
APPLICANT: Wallis, Nicola G.
APPLICANT: Shilling, Lisa K.
APPLICANT: Wang, Min
APPLICANT: Jaworski, Deborah D.
APPLICANT: Ingram, Karen A.
APPLICANT: Yigang Ge, James
APPLICANT: Holmes, David J.
APPLICANT: Zalacain, Magdalena
Throup, John

APPLICANT: Batswag, Sanjoy
TITLE OF INVENTION: Histidine Kinase
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Street
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/081,686
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/048,339
FILING DATE: 30-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Palk, Stephen T
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GM10006
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1458 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

Query Match	17.2%	Score 32.4	DB 3	Length 1458
Best Local Similarity	26.6%	Pred. No. 1.8		
Matches	28	Conservative	29	Mismatches 41, Indels 0, Gaps 0

QY	2	AAAAAUGAUCUUCGUCUGUAAAUACAUAUUUGAGAGUAAUUAUUACAUAAGUCUC	61
		: : : : : : : : : : : : : : : : : : : :	
Db	121	AAAAATCTAAATATATTAACCAATCTTTTAAAGAGTTTGCATATCTAGTGTGT	180
		: : : : : : : : : : : : : : : : : : : :	
QY	62	AUUUUUGAUUUUGAGUUGCUAUUUUACUUUACGUUCC	99
		: : : : : : : : : : : : : : : : : : : :	
Db	181	GGCTCTAGTCATTCATATAGCTATTATTATTAACCTTTCC	218
		: : : : : : : : : : : : : : : : : : : :	

RESULT 12
 US-09-081-686-3
 ; Sequence 3, Application US/09081686
 ; Patent No. 6162619
 ;
 ; GENERAL INFORMATION:
 ; APPLICANT: Wallis, Nicola G.
 ; APPLICANT: Shilling, Lita K.
 ; APPLICANT: Wang, Min
 ; APPLICANT: Jaworski, Deborah D.
 ; APPLICANT: Ingramham, Karen A.
 ; APPLICANT: Yigong Ge, James A.
 ; APPLICANT: Holmes, David J.
 ; APPLICANT: Zalcstein, Magdalena
 ; APPLICANT: Throup, John
 ; APPLICANT: Biswas, Sanjoy
 ; TITLE OF INVENTION: Hseltidine Kinase
 ;
 ; NUMBER OF SEQUENCES: 9
 ;
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSSEE: Dechert, Price & Rhoads
 ; STREET: 4000 Bell Atlantic Tower, 1717 Arch Street
 ; CITY: Philadelphia
 ; STATE: PA


```

; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081.686
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,339
; FILING DATE: 30-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-09-081-686-3

```

```

Query Match      17.2% Score 32.4; DB 3; Length 2279;
Best Local Similarity 28.6%; Pred. No. 2.2;
Matches 28; Conservative 29; Mismatches 41; Indels 0; Gaps 0;

QY      2  AAAAUGUGAUCUCUGUAAAUAACAUAUUUGAGAGUUAUAUAUACAAGUAGUCU 61
DB      1322 AAAATCTCTAATATTATTAACCAATCTTTTAAGAAGTTTGCAATTCTAGTGAGTGT 1381

QY      62 AUUUUGUAUUUAGUUGUAUUUAGCUUUUAGCUUUC 99
DB      1382 GGTCTAGTCATTCATATAGCTATTATTATTTGACCTTTC 1419

```

```

RESULT 13
US-09-543-681A-1309
; Sequence 1309, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1309
; LENGTH: 402
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-1309

```

```

Query Match      17.1% Score 32.2; DB 3; Length 402;
Best Local Similarity 32.6%; Pred. No. 1.3;
Matches 46; Conservative 27; Mismatches 68; Indels 0; Gaps 0;

QY      37 AGGUUAUAUAUAACAAGUAGUUAUUUUGUAUUUAGUUGUAUUUAGCUUUUAGCU 96
DB      138 AAGCTCATCAATTCATTTGCTGCTATATTAAGATCTCGTTACTATTTTCCCTAACGT 197

QY      97 UCCAGAGUCCUGAGUCCAGCCCAUAUUCAGAGAGCCUUCUGCGAUUUUUCAGA 156

```

```

DB      198 TCCACTAATTAATCAATTAACCATATTCATAATCTATACAGTGTGTCGA 257
QY      157 UUAGUAGUGCAAAACCUAA 177
DB      258 TGAGATTCCGATATATCTAA 278

```

```

RESULT 14
US-09-543-681A-1449/C
; Sequence 1449, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1449
; LENGTH: 1407
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-1449

```

```

Query Match      17.1% Score 32.2; DB 3; Length 1407;
Best Local Similarity 32.6%; Pred. No. 2.1;
Matches 46; Conservative 27; Mismatches 68; Indels 0; Gaps 0;

QY      37 AGGUUAUAUAUAACAAGUAGUUAUUUUGUAUUUAGUUGUAUUUAGCUUUUAGCU 96
DB      840 AAGCTCATCAATTCATTTGCTGCTATATTAAGATCTCGTTACTATTTTCCCTAACGT 781

QY      97 UCCAGAGUCCUGAGUCCAGCCCAUAUUCAGAGAGCCUUCUGCGAUUUUUCAGA 156
DB      780 TCCACTAATTAATCAATTAACATATTCATATTAATCTATACAGTGTGTCGA 721

QY      157 UUAGUAGUGCAAAACCUAA 177
DB      720 TGAGATTCCGATATATCTAA 700

```

```

RESULT 15
US-09-949-016-12298/C
; Sequence 12298, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949.016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12298
; LENGTH: 115963
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(115963)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12298

```


US-10-088-750B-3
; Sequence 3, Application US/10088750B
; Publication No. US20040166486A1
; GENERAL INFORMATION:
; APPLICANT: NAKASHIMA, Nobuhiko
; APPLICANT: KANAMORI, Yasushi
; TITLE OF INVENTION: A Novel Higher-Order Structure With Promoting Translation
; FILE REFERENCE: 3190-015
; CURRENT APPLICATION NUMBER: US/10/088,750B
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: JP P2001-016746
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/JP01/00641
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 201
; TYPE: RNA
; ORGANISM: Drosophila C Virus
US-10-088-750B-3

Query Match 65.7%; Score 123.6; DB 7; Length 201;
Best Local Similarity 79.0%; Pred. No. 1.1e-25;
Matches 147; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 3 AAAAUGUGAUCUGUGUAAUACAUAUUUGAGAGUAAUAAUUAUACAAGUUGUGUA 62
DB 4 AAGAGUGUACUUCUUCUUCUUAUACAUAUUUGAGAGUAAUAAUACAAGUUGUGUA 63
QY 63 UUUUUGAUAUUUGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 122
DB 64 UCUUAUAUUUGAGUUAUAUAUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 123
QY 123 AAUAUCCGAGAACCCUUCUUCGCGUUUUUUCAGUAUUAGUUGGCAAAAACCUAAGAAU 182
DB 124 AAUAUCCGAGAACCCUUCUUCGCGUUUAUAUUAUUAUUAUUAUUAUUAUUAUUA 183
QY 183 UUAACU 188
DB 184 UAACU 189

RESULT 3
US-10-088-750B-6
; Sequence 6, Application US/10088750B
; Publication No. US20040166486A1
; GENERAL INFORMATION:
; APPLICANT: NAKASHIMA, Nobuhiko
; APPLICANT: KANAMORI, Yasushi
; TITLE OF INVENTION: A Novel Higher-Order Structure With Promoting Translation
; FILE REFERENCE: 3190-015
; CURRENT APPLICATION NUMBER: US/10/088,750B
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: JP P2001-016746
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/JP01/00641
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 202
; TYPE: RNA
; ORGANISM: Black Queen-Cell Virus
US-10-088-750B-6

Query Match 33.7%; Score 63.4; DB 7; Length 202;
Best Local Similarity 65.9%; Pred. No. 3.2e-08;
Matches 122; Conservative 0; Mismatches 61; Indels 2; Gaps 2;
QY 1 CAAAAUUGAUCUGUGUAAUACAUAUUUGAGAGUAAUAAUUAUACAAGUUGUG 59

DB 2 CAAUUGUGAUCUGUGCGAGGCAAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 61
QY 60 CUUUUUUGAUAUUUGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 118
DB 62 CUUUUUUGAUAUCCGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 121
QY 119 CCAUAUAUCCAGAACCCUUCUUCGCGUUUUUUCAGUAUUAGUUGGCAAAAACCUAAG 178
DB 122 CUUAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 181
QY 179 AAUU 183
DB 182 AAUU 186

RESULT 4
US-10-088-750B-2
; Sequence 2, Application US/10088750B
; Publication No. US20040166486A1
; GENERAL INFORMATION:
; APPLICANT: NAKASHIMA, Nobuhiko
; APPLICANT: KANAMORI, Yasushi
; TITLE OF INVENTION: A Novel Higher-Order Structure With Promoting Translation
; FILE REFERENCE: 3190-015
; CURRENT APPLICATION NUMBER: US/10/088,750B
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: JP P2001-016746
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/JP01/00641
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 199
; TYPE: RNA
; ORGANISM: Himerobol P Virus
US-10-088-750B-2

Query Match 24.7%; Score 46.4; DB 7; Length 199;
Best Local Similarity 58.1%; Pred. No. 0.0027;
Matches 100; Conservative 0; Mismatches 71; Indels 1; Gaps 1;
QY 15 UGCUUGAAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 74
DB 15 UGAUUAAGAGUAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 74
QY 75 GGUUAGCUUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 134
DB 75 CUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 134
QY 135 GCCUUCUUCGCGUUUUUUCAGUAUUAGUUGGCAAAAACCUAAGAAUUAUUAUUA 186
DB 135 A-CCUAGUGAGCCUUGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 185

RESULT 5
US-09-925-065A-778298
; Sequence 778298, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092


```

; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luehly, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madsen, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKiel, Christine L.
; APPLICANT: Miller, Philip W.
; APPLICANT: Padmavathi, Manohikanti
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Tennesen, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyun
; APPLICANT: Xin, Zhanguo
; APPLICANT: Xu, Nanfei
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 25
; LENGTH: 6440
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1552)..(6087)
; OTHER INFORMATION:
US-10-310-154-25

```

Query Match 19.4%; Score 36.4; DB 6; Length 6440;

Best Local Similarity 31.9%; Pred. No. 9.1; Matches 30; Conservative 28; Mismatches 36; Indels 0; Gaps 0;

```

Qy 5 AAUGAUCUCUGCUUUAUACAUAUUUUGAGAGUUAUAUUACAAGUGCUAUU 64
Db 128 AATTGATCTGGCTCGACATTAGTATTCAATGAGAGATTAAGTTTGTAGTCTT 69
Qy 65 UUUUGAUUUAGUUGCUAUUUUAGCUUUGUC 98
Db 68 TTTTCTATCTTTAGCTAAGTTCCTAACGCTC 35

```

RESULT 9

```

; Sequence 76, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 76
; LENGTH: 6440
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-732-923-76

```

Query Match 19.4%; Score 36.4; DB 9; Length 6440;
Best Local Similarity 31.9%; Pred. No. 9.1;

```

Matches 30; Conservative 28; Mismatches 36; Indels 0; Gaps 0;
Qy 5 AAUGAUCUCUGCUUUAUACAUAUUUUGAGAGUUAUAUUACAAGUGCUAUU 64
Db 128 AATTGATCTGGCTCGACATTAGTATTCAATGAGAGATTAAGTTTGTAGTCTT 69
Qy 65 UUUUGAUUUAGUUGCUAUUUUAGCUUUGUC 98
Db 68 TTTTCTATCTTTAGCTAAGTTCCTAACGCTC 35

```

RESULT 10

```

; Sequence 53, Application US/10078770
; Publication No. US20030003471A1
; GENERAL INFORMATION:
; APPLICANT: Farnod, Omolayo O.
; APPLICANT: Forge, Charlie
; APPLICANT: Miao, Guo-Hua
; TITLE OF INVENTION: CDNAS Encoding Polypeptides
; FILE REFERENCE: BB-1365 US NA
; CURRENT APPLICATION NUMBER: US/10/078,770
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 09/614,188
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,400
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/153,534
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 60/161,223
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/159,878
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/157,401
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/143,419
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,409
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 53
; LENGTH: 2480
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-078-770-53

```

Query Match 19.0%; Score 35.8; DB 5; Length 2480;

Best Local Similarity 36.9%; Pred. No. 9.2; Matches 38; Conservative 23; Mismatches 42; Indels 0; Gaps 0;

```

Qy 3 AAAAUGAUCUCUGCUUUAUACAUAUUUUGAGAGUUAUAUUACAAGUGCUA 62
Db 2218 AAGATTGCTTAACCTTGTATATATACAGAGATGGAAACAAATATATGTGA 2277
Qy 63 UUUUGAUUUAGUUGCUAUUUUAGCUUUGUCGAGAU 105
Db 2278 TTTTCTATCTTTAGTGTAGTGTGATGATTTGGTCAATTG 2320

```

RESULT 11

```

; Sequence 96143, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
US-10-437-963-96143

```

Query Match 19.4%; Score 36.4; DB 9; Length 6440;
Best Local Similarity 31.9%; Pred. No. 9.1;


```

; APPLICANT: OLEK, Alexander
; APPLICANT: PIERENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 312
; LENGTH: 7312
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically created genomic DNA (Homo sapiens)
US-10-221-613-312

Query Match          18.9%; Score 35.6; DB 7; Length 7312;
Best Local Similarity 25.6%; Pred. No. 16;
Matches 23; Conservative 33; Mismatches 34; Indels 0; Gaps 0;

Y 12 UCUCGCGUAAUACAUAUUUGAGGUGUAAUAAAUUACAAGUGGCUAUUUUGAU 71
D 6367 TTTATTATTAATTGAATTTTGAGATGGGATTAATGTAATTAATTATTTTGA 6426
Y 72 UUAGUUAGCUAUUUAGCUUUUAGUCCAG 101
D 6427 TAAGTTTGTATTGTATTATTAGTTGAG 6456

RESULT 14
US-10-088-750B-1
; Sequence 1, Application US/10088750B
; Publication No. US20040166486A1
; GENERAL INFORMATION:
; APPLICANT: NAKASHIMA, Nobuhiko
; APPLICANT: KANAMORI, Yasuhide
; TITLE OF INVENTION: A Novel Higher-Order Structure with Promoting Translation
; FILE REFERENCE: 3190-015
; CURRENT APPLICATION NUMBER: US/10/088,750B
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: JP P2001-016746
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/JP01/00641
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 200
; TYPE: RNA
; ORGANISM: Plautia stali Intestine Virus
US-10-088-750B-1

Query Match          18.4%; Score 34.6; DB 7; Length 200;
Best Local Similarity 56.6%; Pred. No. 7.3;
Matches 64; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Y 28 AAUUUUGAGGUAUUAAAUUACAAGUAGGCUAUUUUUUGAUUUUAGUUUAGCUAUUA 87
D 29 AAUUUUGAGGUAUUAAAUUAGUUUUUUGCUAAGUCUUGAGAGGUCUUGUAUUUUUA 88

```



```

; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47569
; LENGTH: 1131
; TYPE: DNA
; ORGANISM: Bovine 19866881500383
; US-10-750-185-47569
```

```
Query Match      17.7%; Score 33.2; DB 6; Length 1131;
Best Local Similarity 27.5%; Pred. No. 0.62;
Matches 28; Conservative 31; Mismatches 43; Indels 0; Gaps 0;
```

```

QY      7 UGUAUCUCUCUGUAAUACAUAUUUGAGAAGUUAUAAUUAACAAGUAGUCUAUUU 66
DB      1102 TGTGTTTATTTTTCACAAATGCTGAGTTAAATGAATCTTGACTGTGTTCTTT 1043
QY      67 UGUUUUUGUAGUAGUUAUUUGCUUUAAGUCCAGAGAGCCU 108
DB      1042 TATATTAGGTGAGATCATATGTTTACCTTTAAAAAGCCT 1001
```

```

RESULT 3
US-10-873-528-168
; Sequence 168, Application US/10873528
; Publication No. US20050276814A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Phillip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/10/873,528
; CURRENT FILING DATE: 2004-06-23
; PRIOR APPLICATION NUMBER: US/09/769,787
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 168
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; US-10-873-528-168
```

```

Query Match      16.4%; Score 30.8; DB 6; Length 1332;
Best Local Similarity 28.6%; Pred. No. 3.7;
Matches 28; Conservative 28; Mismatches 42; Indels 0; Gaps 0;

QY      2 AAAAUGUAGUCUCUGUAAUACAUAUUUGAGAAGUUAUAAUUAACAAGUAGUCU 61
DB      7 AAAAATCTTAATATTAACCAAGCTTTTAAAGAGTTTGCAATTTAGAGTGGTGT 66
QY      62 AUUUUUUUAUUAGUAGUUAUUUUAAGCUUUAAGCUUCC 99
DB      67 GGTCTAGTCATTCATATAGCATTAATTATTGACCTTTC 104
```

```

RESULT 4
US-10-873-528-213
```

```

; Sequence 213, Application US/10873528
; Publication No. US20050276814A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Phillip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/10/873,528
; CURRENT FILING DATE: 2004-06-23
; PRIOR APPLICATION NUMBER: US/09/769,787
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 213
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; US-10-873-528-213
```

```
Query Match      16.4%; Score 30.8; DB 6; Length 1332;
Best Local Similarity 28.6%; Pred. No. 3.7;
Matches 28; Conservative 28; Mismatches 42; Indels 0; Gaps 0;
```

```

QY      2 AAAAUGUAGUCUCUGUAAUACAUAUUUGAGAAGUUAUAAUUAACAAGUAGUCU 61
DB      7 AAAAATCTTAATATTAACCAAGCTTTTAAAGAGTTTGCAATTTAGAGTGGTGT 66
QY      62 AUUUUUUUAUUAGUAGUUAUUUUAAGCUUUAAGCUUCC 99
DB      67 GGTCTAGTCATTCATATAGCATTAATTATTGACCTTTC 104
```

```

RESULT 5
US-11-121-086-43/C
; Sequence 43, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 0918.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 43
; LENGTH: 175416
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-121-086-43
```

```

Query Match      16.4%; Score 30.8; DB 7; Length 175416;
Best Local Similarity 27.8%; Pred. No. 28;
Matches 25; Conservative 28; Mismatches 37; Indels 0; Gaps 0;

QY      3 AAAAUGUAGUCUCUGUAAUACAUAUUUGAGAAGUUAUAAUUAACAAGUAGUCU 62
DB      60779 AAACGTGACCTAGATAGGCAATGCTGTGAGAAAGCTTTGCCAATAGTGATTTATT 60720
QY      63 UUUUUUUAUUAGUAGUUAUUUUAAGCUUUAAGCUU 92
DB      60719 TTTATTATTAATTATTAATTATTATTTAT 60690
```

```

RESULT 6
US-10-995-561-13487/C
```



```
; Sequence 13487, Application US/1095561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13487
; LENGTH: 41309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13487

Query Match      16.3%; Score 30.6; DB 6; Length 41309;
Best Local Similarity 29.1%; Pred. No. 18;
Matches 34; Conservative 29; Mismatches 54; Indels 0; Gaps 0;

QY      20 GUAAAUACAUAUUUGAGAGUUAUAAUUAACAAGUAGCUAUUUUUUGUAUUUAGGUUA 79
DB      4801 GCATATATGCTTTATGATTTACTTAATACATATAATATATTTATTTAATATG 4742
QY      80 GCUAUUAGCUUUAUAGUCCAGAGUCCUAGUGGACCCCAUAUCCAGAAAC 136
DB      4741 AATATTATATGTTATATGCTTATTTTCGGGAGGAAATGGAAGATATCTCGAAGAC 4685

RESULT 7
US-11-121-086-24/c
; Sequence 24, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24
; LENGTH: 120096
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-24

Query Match      16.3%; Score 30.6; DB 7; Length 120096;
Best Local Similarity 27.7%; Pred. No. 28;
Matches 39; Conservative 33; Mismatches 69; Indels 0; Gaps 0;

QY      10 GAUCUCUCUUAUAUACAUAUUUGAGAGUUAUAAUUAACAAGUAGCUAUUUUUUGU 69
DB      84273 GTTTAGCTTGCGGTAATGAGTTTGTTCATTTTCAGTTAAGAGTAGTTAGTTCTT 84214
QY      70 AUUUAAGUUAUUAUUUUGAGUUAUAGUCCAGAGUCCUAGUGGACCCCAUAUUAUCC 129
DB      84213 GTACATGTTAATACAAACCCCTTTACACAGATGATTCGCGAATCTTTTCATCTGT 84154
QY      130 AGAAGACCCUUCUUCGCGUU 150
DB      84153 AGGATGCTTCTTTAATGTT 84133

RESULT 8
US-11-112-908-65/c
; Sequence 65, Application US/11112908
; Publication No. US2005026059A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 65
; LENGTH: 173115
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-65

Query Match      16.2%; Score 30.4; DB 7; Length 173115;
Best Local Similarity 31.9%; Pred. No. 37;
Matches 46; Conservative 27; Mismatches 71; Indels 0; Gaps 0;

QY      29 AUUUGAGAGUUAUUAUUAUUAACAAGUAGCUAUUUUUUGUAUUUAGGUAGCUAUUUG 88
DB      143289 ATTATTTGGGGGAAATATTTATAGATATATGCTAGGGATGCAAAATATTTGACATTTTATG 143230
QY      89 CUUUAAGUCCAGAGUAGUCCUAGUGGACCCCAUAUUAUCCAGAGGCCUUCUCGCGU 148
DB      143229 TTTTCAGTTACGCTACGTCGCAATGTCCTAGAGAGATATTAATCTGCTGTGCT 143170
QY      149 UUUUCAGAUUAGUAGUUGCAAAA 172
DB      143169 ATGAAATATATCATAGTGTGTTAA 143146

RESULT 9
US-11-102-978-3
; Sequence 3, Application US/11102978
; Publication No. US20050250142A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Technology Transfer Office
; APPLICANT: University of Utah Research Foundation
; TITLE OF INVENTION: Diagnosis and Treatment of Herpes Simplex Virus Disease
; FILE REFERENCE: 0274-5537,1US
; CURRENT APPLICATION NUMBER: US/11/102,978
; PRIOR APPLICATION NUMBER: PCT/US2003/033152
; PRIOR FILING DATE: 2003-10-18
; PRIOR APPLICATION NUMBER: 60/419,576
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 340000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (56948)..(57115)
; OTHER INFORMATION: C21orf34 exon
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8006)..(81089)
; OTHER INFORMATION: Gene VDAC2P, voltage-dependent anion channel isoform 2 pseudogene
; FEATURE:
; NAME/KEY: exon
; LOCATION: (167308)..(167438)
; OTHER INFORMATION: C21orf34 exon
```



```
; FEATURE:
; NAME/KEY: exon
; LOCATION: (216732)..(216833)
; OTHER INFORMATION: C21orf34 exon
US-11-102-978-3

Query Match      16.2%; Score 30.4; DB 7; Length 340000;
Best Local Similarity 26.0%; Pred. No. 49;
Matches 27; Conservative 31; Mismatches 46; Indels 0; Gaps 0;

OY 3 AAAAUGAGUCUCUCUGUAAUACAUAUUUGAGAGGUAAUAAUUAACAAGUAGUCUA 62
Db 113138 AAATTCACATTTTATGACACGATTAATCTCAGATTGACAAATAATAGTGGTACTT 113197
OY 63 UUUUUUAUUUAGUUAAGUUAUUUAGCUUUACGUUCCAGAUCC 106
Db 113198 TTTTCTTTTATTTTATTTTATTTATTAATCTTAAGTTCGGGGTAC 113241

RESULT 10
US-10-750-185-32825
; Sequence 32825, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750.185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32825
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Bovine 19866881322397
US-10-750-185-32825

Query Match      16.1%; Score 30.2; DB 6; Length 723;
Best Local Similarity 28.6%; Pred. No. 4.4;
Matches 26; Conservative 27; Mismatches 38; Indels 0; Gaps 0;

OY 10 GAUCUUCUCUCUUAUAUACAUAUUUGAGAGGUAAUAAUUAACAAGUAGUCUAUUUUU 69
Db 577 GATTTTCTTGCTGATGGAGACTTTTAAGACTTAAGTCACTTCAAGTACACATACAGTACT 636
OY 70 AUUAGGUUAGCUUAUUUAGCUUUACGUUCCA 100
Db 637 GTTAACCATATTAATTCATGCTTTTACATTACA 667

RESULT 11
US-10-750-185-47264
; Sequence 47264, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750.185
```

```
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47264
; LENGTH: 1375
; TYPE: DNA
; ORGANISM: Bovine 19866880820089
US-10-750-185-47264

Query Match      16.1%; Score 30.2; DB 6; Length 1375;
Best Local Similarity 31.3%; Pred. No. 5.7;
Matches 26; Conservative 24; Mismatches 33; Indels 0; Gaps 0;

OY 5 LAUCUGAUCUCUCUGUAAUACAUAUUUGAGAGGUAAUAAUUAACAAGUAGUCUAUU 64
Db 1198 AAGGTGATCTAGCAAAATTAATATTTTAGACATTTACTCATGTGGAAGAAGTGGCTTTT 1257
OY 65 UUUUUUAUUUAGUUAAGCUUAUUUA 87
Db 1258 TTTTCTTTTCAATAGCTATGTA 1280

RESULT 12
US-11-194-246-373
; Sequence 373, Application US/11194246
; Publication No. US20050272089A1
; GENERAL INFORMATION:
; APPLICANT: Mott, John
; APPLICANT: Trepo, Catherine
; APPLICANT: Arvidson, Staffan
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND METI
; FILE REFERENCE: 00592.US1 (MAR 266.05920101)
; CURRENT APPLICATION NUMBER: US/11/194.246
; CURRENT FILING DATE: 2005-08-01
; PRIOR APPLICATION NUMBER: US/10/274.586
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/345,438
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 621
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 373
; LENGTH: 2130
; TYPE: DNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Sequence of the rpl21 gene and flanking regions.
US-11-194-246-373

Query Match      16.1%; Score 30.2; DB 7; Length 2130;
Best Local Similarity 31.8%; Pred. No. 6.9;
Matches 34; Conservative 25; Mismatches 48; Indels 0; Gaps 0;

OY 22 AAAUACAUAUUUGAGAGGUAAUAAUUAACAAGUAGUCUAUUUUUUAUUUAGUUAUC 81
Db 1865 AATTAATGTACGTAATTAATTACACAACAATCTTGCTAATGATGATTAATGAGTTGTG 1924
OY 82 UAUUAGCUUUUAGCUUCCAGAUCCCUUAUGAGACCCCAUAUAC 128
Db 1925 CCATAGCCTTTATGCCGATGCGTACTCTCGCAAGTACAGAAAGATC 1971

RESULT 13
US-11-194-246-129
; Sequence 129, Application US/11194246
; Publication No. US20050272089A1
; GENERAL INFORMATION:
; APPLICANT: Mott, John
; APPLICANT: Trepo, Catherine
; APPLICANT: Arvidson, Staffan
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND METI
```



```

? TITLE OR INVENTION: USE
?
? FILE REFERENCE: 00592.US1 (MAR 268.05920101)
?
? CURRENT APPLICATION NUMBER: US/11/194,246
?
? CURRENT FILING DATE: 2005-08-01
?
? PRIOR APPLICATION NUMBER: US/10/274,586
?
? PRIOR FILING DATE: 2002-10-21
?
? PRIOR APPLICATION NUMBER: US 60/345,438
?
? PRIOR FILING DATE: 2001-10-19
?
? NUMBER OF SEQ ID NOS: 621
?
? SOFTWARE: PatentIn version 3.0
?
? SEQ ID NO 129
?
? LENGTH: 2255
?
? TYPE: DNA
?
? ORGANISM: ARTIFICIAL
?
?
? FEATURE:
?
? OTHER INFORMATION: Sequence of the L27 coding sequence and flanking regions
?
? US-11-194-246-129

```

Query March	16.1%	Score 30.2;	DB 7;	Length 225;
Best Local Similarity	31.8%	Pred. No. 7;		
Matches 34; Conservative	25;	Mismatches 48;	Indels 0;	Gaps 0;

[illegible]

Qy 82 UAUUAGCUUUA CGUUU CCAAGAUCCUAGUGG CAGCCCC CACAUAUUC 128
: |:::| |:::| |:::| |:::|
Db 1925 CCATAGCCTTATG CCGAGTGC TGACTTCTTCG CCAAGTACAAGA CTC 1971

```

RESULT 14
US-10-750-185-57886/c
, Sequence 57886, Application US/10750185
, Publication NO. US2005026063A1
, GENERAL INFORMATION:
, APPLICANT: MMI GENOMICS, INC.
, APPLICANT: DENISE, Sue K.
, APPLICANT: KERR, Richard
, APPLICANT: ROSENFELD, David
, APPLICANT: HOLM, Tom
, APPLICANT: BATES, Stephen
, APPLICANT: PANTIN, Dennis
, TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
, FILE REFERENCE: MM1100-2
, CURRENT APPLICATION NUMBER: US/10/750,185
, CURRENT FILING DATE: 2003-12-31
, PRIOR APPLICATION NUMBER: US 60/437,482
, PRIOR FILING DATE: 2002-12-31
, NUMBER OF SEQ ID NOS: 64922
, SOFTWARE: PatentIn version 3.1
, SEQ ID NO 57886
, LENGTH: 2122
, TYPE: DNA
, ORGANISM: Bovine
US-10-750-185-57886 19866880700652

```

Query Match	15.6%	Score 29.4	DB 6	Length 2122
Similarity	25.3%	Pred. No. 12		
Best Local				
Matches	22	Conservative 29	Mismatches 36	Indels 0
				Gaps 0

Oy 11 AUCCTGCTUGAAAUACAUAUUTUGAGAGCGUUAAUAAAUTACAAGAUGAGCUAUAUUTUGUA 70
| : | : : : | : : : | : : : | : : : | : : :
Db 117 ATCCGTTCATCTTCCATTCTTTCAAGTGTCCTTAATAATACAAATGCTTGTATAGTATG 56

```
Oy      71 UUUAGGCUUAGCUAUUUAGCUUUUACGU  97
      :  ::  :  :  :  :  :  :
Db      57 TCCTAGTTACGTACCTACTTTTAATT  31
```

RESULT 15
US-10-750-185-41669
; Sequence 41669, Application US/10750185

```

: Publication No. US20050260603A1
:
: GENERAL INFORMATION:
: APPLICANT: MMT GENOMICS, INC.
: APPLICANT: DENISE, Sue K.
: APPLICANT: KERR, Richard
: APPLICANT: ROSENFELD, David
: APPLICANT: HOLM, Tom
: APPLICANT: BATES, Stephen
: APPLICANT: PARTIN, Dennis
: TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
:
: FILE REFERENCE: MM1100-2
: CURRENT APPLICATION NUMBER: US/10/750,185
: CURRENT FILING DATE: 2003-12-31
: PRIOR APPLICATION NUMBER: US 60/437,482
: PRIOR FILING DATE: 2002-12-31
:
: NUMBER OF SEQ ID NOS: 64922
:
: SOFTWARE: PatentIn version 3.1
:
: SEQ ID NO 41669
:
: LENGTH: 1228
:
: TYPE: DNA
:
: ORGANISM: Bovine
:
: US-10-750-185-41669

```

Query Match	15.5%	Score 29.2;	DB 6;	Length 1228;
Best Local Similarity	35.8%	Pred. No. 11;		
Matches 38;	Conservative 20;	Mismatches 48;	Indels 0;	Gaps 0;

Dy 30 TTUUGAGAGGUTAAUUAAAUTUACAAGUAGUCUAUUUUUUUGAUUUUAGGUUAGCUAUUUAGC 89
::: ||| : | : | : :::: : : |
Db 322 TTTTCATTGGCTCAACCATGCATAATTAGCATTAATTTTTAGCTTCAAAGACTAAAGCAAAC 38Db

```

Qy      90  UUUACGUUCCAGGAUGCCUAGUGGCAGCCCCCAUAUUCAGGAAG 135
          :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      382  TATACATGAAAGAAATCCAGAGGCAACCAATTAAATTCAGGAAG 427

```

Search completed: December 22, 2005, 16:41:06
Job time : 174.55 secs

Search completed: December 22, 2005, 16:41:06
Job time : 174.55 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 04:46:32 ; Search time 1467.3 Seconds
(without alignments)
7205.664 Million cell updates/sec

Title: US-10-088-750C-5

Perfect score: 186
Sequence: 1 uugacuaugauugauugcu.....uugacuaugauugacuc 186

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_str:*
11: gb_ey:*
12: gb_un:*
13: gb_vl:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	186	100.0	186 6 BD173515	BD173515 Novel ter
2	186	100.0	186 6 BD177019	BD177019 Novel ter
3	175	94.1	9010 13 AF178440	AF178440 Triatoma
4	36.8	19.8	9812 13 AF536531	AF536531 Aphid let
5	36.8	19.8	9829 6 AX763949	AX763949 Sequence
6	36.4	19.6	189 6 BD173513	BD173513 Novel ter
7	36.4	19.6	189 6 BD177017	BD177017 Novel ter
8	36.4	19.6	1345 6 AR496112	AR496112 Sequence
9	36.4	19.6	1345 6 AR511394	AR511394 Sequence
10	36.4	19.6	9264 13 AF014388	AF014388 Drosophila
11	36	19.4	128544 8 AF238378	AF238378 Homo sapi
12	35.8	19.2	636 10 BV432331	BV432331 S237P6298
13	35	18.8	239102 14 AC137051	AC137051 Rattus no
14	35	18.8	267802 14 AC137461	AC137461 Rattus no
15	34.8	18.7	126572 5 BX784037	BX784037 Zebrafish
16	34.8	18.7	140919 8 AC110747	AC110747 Homo sapi
17	34.8	18.7	168355 5 BX649631	BX649631 Zebrafish
18	34.6	18.6	157381 8 AL360267	AL360267 Human DNA

c 19	34.6	18.6	175331 14 AC079275	AC079275 Mus muscu
c 20	34.6	18.6	199573 9 AC154124	AC154124 Mus muscu
c 21	34.6	18.6	216375 14 AC119131	AC119131 Rattus no
c 22	34.6	18.6	248936 14 AC135200	AC135200 Rattus no
c 23	34.6	18.6	252978 14 AC097297	AC097297 Rattus no
c 24	34.4	18.5	125970 14 AC136841	AC136841 Medicago
c 25	34.4	18.5	226770 9 AC159890	AC159890 Mus muscu
c 26	34.4	18.5	243345 14 CR792426	CR792426 Dario rer
c 27	34.2	18.4	235547 14 AC163001	AC163001 Bos tauru
c 28	33.6	18.1	188 6 BD173514	BD173514 Novel ter
c 29	33.6	18.1	188 6 BD177018	BD177018 Novel ter
c 30	33.6	18.1	9185 13 AF218039	AF218039 Crickel p
c 31	33.6	18.1	110000 14 AC094850_2	Continuation (3 of
c 32	33.6	18.1	229829 14 AC132064	AC132064 Rattus no
c 33	33.6	18.1	234284 14 AC139280	AC139280 Rattus no
c 34	33.6	18.1	267693 14 AC111954	AC111954 Rattus no
c 35	33.6	18.1	267693 14 AC120667	AC120667 Rattus no
c 36	33.4	18.0	154434 14 AC135887	AC135887 Rattus no
c 37	33.4	18.0	209017 14 AC134209	AC134209 Rattus no
c 38	33.4	18.0	222315 9 AL593847	AL593847 Mouse DNA
c 39	33.2	17.8	110000 1 CP000003_07	Continuation (8 of
c 40	33.2	17.8	172778 5 CR788302	CR788302 Zebrafish
c 41	33	17.7	96469 14 AP008180	AP008180 Locust cor
c 42	33	17.7	191069 14 AC159981	AC159981 Mus muscu
c 43	33	17.7	201617 9 AC109175	AC109175 Mus muscu
c 44	33	17.7	203115 14 AC124559	AC124559 Mus muscu
c 45	33	17.7	242003 14 AC133103	AC133103 Mus muscu

ALIGNMENTS

RESULT 1	BD173515	186 bp	RNA	linear	PAT 18-FEB-2003
LOCUS	BD173515				
DEFINITION	Novel tertiary structure having ability to accelerate translation				
ACCESSION	BD173515				
VERSION	BD173515.1	GI:26414846			
KEYWORDS	WO 02061080-A/5.				
SOURCE	Triatoma virus				
ORGANISM	Triatoma virus				
REFERENCE	Viruses; ssRNA positive-strand viruses, no DNA stage; Dicistroviridae; Crispavirus.				
AUTHORS	1 (bases 1 to 186)				
TITLE	Nakashima,N. and Kanamori,Y.				
JOURNAL	Novel tertiary structure having ability to accelerate translation activity				
Patent:	WO 02061080-A 5 08-AUG-2002;				
JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF					
SERICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE					
FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA, YASUSHI KANAMORI					
OS	Triatoma virus				
PN	WO 02061080-A/5				
PD	08-AUG-2002				
PF	31-JAN-2001	WO 2001JP000641			
PR	25-JAN-2001	JP 01P 016746			
PT	NOBUHIKO NAKASHIMA, YASUSHI KANAMORI				
PC	C12N15/11, C12N15/86, C12P21/02				
CC	Novel tertiary structure having ability to accelerate translation activity				
FH	Key	Location/Qualifiers			
FT	Source	1. 186			
FEATURES	Location/Qualifiers				
source	1. 186				
	/organism="Triatoma virus"				
	/mol_type="genomic RNA"				
	/db_xref="taxon:103442"				
ORIGIN					
Query Match	100.0%;	Score 186;	DB 6;	Length 186;	
Best local Similarity	67.7%;	Pred. No. 1.8e-45;			

Matches	126;	Conservative	60;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	UUGACUUAUUGUAGUCUUCUGUAUUAUUAUUGUACUAUAAAGUGAAGUAGUAGUCUA	60						
DB	1	TTGACATATGTGATCTTGCTTGTGTAATAAATCTGTACATAAAGTCGAAAGTATTGCTA	60						
QY	61	UAGUUAAGUGUGCGUCUUAUUUAGGCAUUCUCUGUGGCGUGUGUAGUCCAA	120						
DB	61	TAGTTAAGGTTGGCTTGCTCTATTATAGGCACTTCTCGAGATGGCGGCTTGCACTCCA	120						
QY	121	CAAGAUCCAGGAGCUGUACAGAUUUUCCUAUACCCUGAGUGCGGUGUUGAUCUAAGU	180						
DB	121	CAAGATCCAGGAGCTGTACAGAAATTTCTATACCTCGAGTCGGGTTGGAATCTAAGGT	180						
QY	181	UGACUC	186						
DB	181	TGACTC	186						
RESULT 2									
LOCUS	BD177019	186 bp	RNA	linear	PAT 16-APR-2003				
DEFINITION	Novel translational activity-promoting higher-order structure.								
ACCESSION	BD177019								
VERSION	BD177019.1	GI:30014279							
KEYWORDS	JP 2002306168-A/5.								
SOURCE	Tritatoma virus								
ORGANISM	Tritatoma virus								
REFERENCE	1	(bases 1 to 186)							
AUTHORS	Nakashima,N. and Kanamori,Y.								
TITLE	Novel translational activity-promoting higher-order structure								
JOURNAL	Patent: JP 2002306168-A 5 22-OCT-2002;								
	DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND								
	ENTOMOLOGICAL HIROSHIMA PREFECTURAL INDUSTRIAL TECHNOLOGY PROMOTION								
	ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND								
	FISHERIES								
COMMENT	OS Tritatoma virus								
	PN JP 2002306168-A/5								
	PD 22-OCT-2002								
	PF 25-JAN-2001								
	PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI								
	PC C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02// PC								
	(C12N15/09, C12N1:92), C12N15/00, C12N5/00, (C12N15/00, C12R1:92) CC								
	Novel translational activity-promoting higher-order structure FH								
	Key								
	Location/Qualifiers								
	FT source	1..186							
	Location/Qualifiers	1..186							
FEATURES	Location/Qualifiers	1..186							
source	/organism="Tritatoma virus"								
	/mol_type="genomic RNA"								
	/db_xref="taxon:103442"								
ORIGIN									
Query Match	100.0%; Score 186; DB 6; Length 186;								
Best Local Similarity	67.7%; Pred. No. 1.8e-45;								
Matches 126; Conservative	60; Mismatches 0; Indels 0; Gaps 0;								
QY	1	UUGACUUAUUGUAGUCUUCUGUAUUAUUAUUGUACUAUAAAGUGAAGUAGUAGUCUA	60						
DB	1	TTGACATATGTGATCTTGCTTGTGTAATAAATCTGTACATAAAGTCGAAAGTATTGCTA	60						
QY	61	UAGUUAAGUGUGCGUCUUAUUUAGGCAUUCUCUGUGGCGUGUGUAGUCCAA	120						
DB	61	TAGTTAAGGTTGGCTTGCTCTATTATAGGCACTTCTCGAGATGGCGGCTTGCACTCCA	120						
QY	121	CAAGAUCCAGGAGCUGUACAGAUUUUCCUAUACCCUGAGUGCGGUGUUGAUCUAAGU	180						
DB	121	CAAGATCCAGGAGCTGTACAGAAATTTCTATACCTCGAGTCGGGTTGGAATCTAAGGT	180						
QY	181	UGACUC	186						

DB	181	TGACTC	186						
RESULT 3									
AF178440									
LOCUS	AF178440	9010 bp	RNA	linear	VR1 14-APR-2000				
DEFINITION	Tritatoma virus nonstructural protein precursor and capsid protein								
ACCESSION	AF178440								
VERSION	AF178440								
KEYWORDS	precursor, gene, partial cds.								
SOURCE	AF178440.1	GI:6003484							
ORGANISM	Tritatoma virus								
REFERENCE	1	(bases 1 to 9010)							
AUTHORS	Cibbenet,C., La Torre,J.L., Muscio,O.A., Ugalde,R.A. and								
TITLE	Nucleotide sequence analysis of Tritatoma virus shows that it is a								
JOURNAL	member of a novel group of insect RNA viruses								
PUBMED	J. Gen. Virol. 81 (Pt 4), 1149-1154 (2000)								
REFERENCE	2	(bases 1 to 9010)							
AUTHORS	Scodeller,E.A. and Cibbenet,C.								
TITLE	Submitted (17-AUG-1999) Centro de Virologia Animal, Serrano 669,								
JOURNAL	Capital Federal, Buenos Aires, Argentina								
FEATURES	Location/Qualifiers								
source	1..9010								
	/organism="Tritatoma virus"								
	/mol_type="genomic RNA"								
	/specific_host="Tritatoma infestans"								
	/db_xref="taxon:103442"								
	/note="insect virus"								
	<549..5936								
	/note="mechanism of expression unknown"								
	/codon_start=1								
	/product="nonstructural protein precursor"								
	/protein_id="AF00472.1"								
	/db_xref="GI:6003485"								
	/translation="EFLRKFLIPFLWCVCVCFQGRISLSORPMPENMKILQKFE								
	KRPSPSLISVSELVEKPPVFLIRQYNFLADDKHMIMTSSKONSTIDLYGNYRFE								
	KYPSFPMWVCINCCEKTEYDSERYENIGMFDPIDPCINGIVDRYERLRGWE								
	SKISYTNLSLILCLFSSLSERSLOOMPERSISPEOSBLPLNGLIKPCRPASSY								
	YNNRYNNRKLVLNPTMPKLTCTNIAKSKRISDBKSCNSTITLCCREDITVLRKD								
	ITPLNLFDLKELIAGSGDOEPTSTVSIINFQORSQTYVNFNORSGESFSQIRPSTV								
	EAFCIQHSDLSRISMSSEPFKLVDRYKSRNGBOIYVESQALSVLPNLSLOPL								
	IDQLCSLSSISDLFSMGKVLIOAIVICSVLSAVLYNLYPFAVSGFCLICSLY								
	FAEDVLCYMRETILDYITPOAOSGGIESLMPCTNYPNALFPNFKLSISEFARFAT								
	VPRALOGIDICIGELISKAPEFARISFYKVLVSDVPITQESVPTKSGEVOKYWKVL								
	ESOLVYDEINENLGLSIOGIVMNSHIFUTARPLISRLAIISILMEFNKRPSE								
	STRNRPVITYISGPTGCKSTLSTPLADITIKIAGDTHDKNNMKLITRNSQ								
	ETWDGTGQLCCVDFPGQRIDTSSNPMLIEFETIRANMYRPLMLMELSGKNTFF								
	SSKIVMCSTNVLREDIKTESLNFPIALKRPRDINVRSLKPIDESLMSKSRPDSI								
	YCFTEVDYNGSIVGALSYKELVNRICSLYSRSPADVAQSIDSVTRKEDLSEBRYAO								
	VGLYFGDDSKSVRYLSVAIPSLIESHSDMKRMNNLIBKYPPIPSMGVGVWGL								
	AVGLIDTSPFKTEKESHVYFSBSCBEKVRVYVCESGEODKQRPVYKCESTL								
	DMKDNQPCDELTGSGCYLDNABEVLAKITKULYMYTMGCVRLGHALPTKGVA								
	PFTHPIAAKCLRISKPEAVECFKSPFNRPVWKMLADLNLSPFSYKTNLPCLDT								
	LAPIRVAMSHNDITDLPVASEVSRVISEPALPILMDNGDPRGVYKFSASGRLT								
	QKRTDLIKYKPSQNIYLRSAWEYSLDTQSGCGAPLILRPMCRGKICGHIVALPT								
	GGAGYAVPTIKFPIIDCLSPFAKBDLTTCVNPESISOGSGVQPLNPEITWNDLST								
	PGEFGVAGKARAIIPSPKSOITELSHVIAVPRRTLLISPIKTDGSPMPMHERMR								
	KYGRPLTALDEIVSTGCAALLNDSTLARKIDITYTNIRSGIYDQTMASGIDDETI								
	NSIKRKSPPGFPMVKTSGTGKQIRFNDGFLPDTPLAVELEKVEVIDLAKQSV								
	RYSHVFDALDKERPRAKHTRAFSGCPLEYLAIVCKMYRGIVSLTKCNERHIS								
	VGTNYSDKWDPMFARYLKSQSDGFVAGDPFBGDSQVPIILREIGNVNGIARQPPD								
	KPEDDEVRLVLIQSLMHSIHNGDVMGMGHALPSGHVLTAPYNLSVATLFEPMAFVI								
	LSRRNGTVMGPMSLMSKFFKEGFPVAYDDHICAPKXRYOSFPNMTLEKVLIELG								
	YTTTEKREIDVPIRSIDELIAYIKRSFVYDERQGMIALITDITVLETSWTHRCDDPI								
	EATVENIFPALRELSHSKQEMEKAPVWMLSEVTRIGRTTTFHDADTRAFVLDLSP								
	DEINMPDLTLM"								

CDS

<6109..8715
/notes="mechanism of expression unknown"
/codon_start=1
/product="capsid protein precursor"
/protein_id="AA00473.1"
/db_xref="GI:6003486"
/translation="LAVNNVNMKQMNVNSQDTTFEORSEKVOAGEINSEIERNQIT
TTTPNDPITTEOLIGDSPOBSVSDRHTSIIDLERPOFGSLPMNTSDIN
KEIFSLPLDPAIAMPREKISGTSFSASVPHIQVNAHPQCGRLVLAAPVDLL
PLHNLNLSFVSNVITLPHVQDLSKETVELKIPYSPVQYDLVTKTPMAFLA
HYVAPLNTPSAASLQVNVFAHFEDIKLGFPISAVIAQAGEQLAPISGLNTVAVG
GVADAVGFSPISIGKADPLVIGNGLTGLSALGFSKPLTTIPITIVQSPQVFN
ADGDGAPLSIKGNEVILKTPAGTSDEVALEYLKIPIYSPFKVSTSPKV
LMTSPVHPOIIRNHTVVDVADAGOPTLLAVATGPFKMRGIVYRFRVKTNYSGRO
ITRPPFGYDDVMSDGKIVRDEYVYRVVDLRQTEATLVVPTSLTPKYCAQVNR
SANPKVNEBRDRKVDNTTDOFTGTGLCVSALTPLVSSAVVSTIDVAVKASD
DFEVAVPNTPLMPLVDSITERPSLDGVIAQVGFASAGTRDIRSVYEGKFIQDITG
MSRNLHDEQSPQICIGERILSFSELIKRNSMRVSDKSLIYPVAFEDNPAAYTA
DKLPVTLTPRSGEPITLTISIGAMYAFYRGGLIKIVGVADQKPLVEVLFPMOD
GVLIKANDYSTPCCSSNIYENFVTKGIAEVOYTPYSRVSVSAVLVNAAGNTSPIM
PNWYKITSNSNILLGHSAADPRRFGPLGAPLAIATALRDNFTGSSATVSLPTFS
NFVLSSTSEST"

ORIGIN

Query Match 94.1%; Score 175; DB 13; Length 9010;
Best Local Similarity 67.4%; Pred. No. 2.9e-42;
Matches 126; Conservative 60; Mismatches 0; Indels 1; Gaps 1;

Qy 1 UUGACUUAUGUAGUCUCUUCUUAUUAUAAA-UUCUGUACAUAUAAAGUCGAAAGUAGUUCU 59
Db 5925 TTGACTAATGATCTTCTGTTCTGTAATAAAATTCGTATACATAAAGTCGAAAGTATGCT 5984
Qy 60 AUAUGUAAAGUUGCGCUCUGCUUAUUAAGCAUACUUCUCAGAGAGCGCGUCGACGUCCA 119
Db 5985 ATAGTTAAGGTGGCGCTTGCTTATTAAGCACTTCTCAGAGATGCGCGCTTGACAGTCCA 6044
Qy 120 ACAAGUCCAGGAGGACUGUACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 179
Db 6045 ACAAGATCCAGGAGCTGACGAATTTCTTACTTACTCTGAGTCCGAGTTTGAAATCTAAG 6104
Qy 180 UUGACUC 186
Db 6105 TTGACTC 6111

RESULT 4
AF536531 9812 bp RNA linear VRL 16-DEC-2002

LOCUS AF536531 Aphid lethal paralysis virus, complete genome.
DEFINITION AF536531
ACCESSION AF536531.1 GI:24431318

KEYWORDS
SOURCE
ORGANISM
Aphid lethal paralysis virus
Aphid lethal paralysis virus
Viruses; ssRNA positive-strand viruses, no DNA stage;
Dicistroviridae; Crispavirinae.

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
12466490
2 (bases 1 to 9812)
van Munster, M., van den Heuvel, J.F.J.M. and van der Wilk, F.
Sequence analysis and genomic organization of Aphid lethal
paralysis virus: a new member of the family Dicistroviridae
J. Gen. Virol. 83 (Pt 12), 3131-3138 (2002)

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
12466490
2 (bases 1 to 9812)
van Munster, M., van den Heuvel, J.F.J.M. and van der Wilk, F.
Sequence analysis and genomic organization of Aphid lethal
paralysis virus: a new member of the family Dicistroviridae
J. Gen. Virol. 83 (Pt 12), 3131-3138 (2002)

FEATURES
source
Location/Qualifiers
1..9812
/organism="Aphid lethal paralysis virus"
/mol_type="genomic RNA"

CDS

/db_xref="taxon:209529"
540..6647
/note="encodes putative helicase, protease, and
RNA-dependent RNA polymerase"
/codon_start=1
/product="nonstructural polyprotein"
/protein_id="AA01470.1"
/db_xref="GI:24431319"

CDS

/translation="MSGSTSTLPLKQLQSLPELNLQKLRDAQAHKRPVLAQTLTS
TLENNNTYIMTEPDLIPENTYLRKQTLATITPENESHGVSRRRAQLDGNIAVLT
LYLDVLMQHTKHAQYDCEFLMKALAMAKENHFLIDGSPYSIALRLNLSALVLT
LNKEDMLKFNLPSCYRLINFLKRIINSSVLTGLLPKRYIDECQSAVANNI
LNMEFRYIOLIKOMTNPNIRNAFIMIPAFYTAFTADITSGSESETLQYV
TVYNKLSIAQGLNENHNIYDNNLELINSIMFADTIVASIKLOEVDVQR
KIYALITTYINIVRSTRKMDLQDLYVNVGALGASGILASVLAQVLTSLRAYVAPR
AQIDTSGHLILKLSLTMFCFLPKQPKNTIDEPTRIDRPKASIGLESMSKID
VVNQIYITFEEELIKKRNRRNTDMQDVLDMSEVKYIELARNIKDPAVLVA
SKLYPRGRVRIKECTQLKSPANLNLIRSLPATKQALADEAKISGALKQALREPLIV
MFCGSGMGKMGMSPEMMDMRVFGDIPTDFQKNIYGRVETPMGTYDQEYIYD
DAFGIKDNVLPKNPELFEIIRLGAPVPMVLMASVEKNTFAPKCVLITSLNDRIK
TESLNSPEAQRRIIDFAYNDIDAEYREYVDMGNEKRYLNAVKAARAKIYGTSV
SNNDVYRFTKFSADGRTIETDNTTAQVTEBCKMKASRPNQMDNTYIADYRNP
YEPEDIVEDPIDEAQAIGNAARVAAGNVAFVLTNHLKRVHGDQDITLIALRKTK
RTMENITLDYGLIPRADIMEQSPMDYHLVYARQIKQLDASTGIMETLAGFQXW
OFYQAVLVGVSTYMCVSPALGDYKGVLPXYIDNDQDAQVLSANACFDK
GNNCKTCKNAANMQLVNMNNSCLCYKMEKMKMINKYAISLYGDORSNHEL
VSDMAEIIINQADDCCTCCNDPELIDALSTYARHGTNCVCLLRFQGRCTADL
YDMTKQISNLPSTIKPELRILDSHARVYDNASRRANSTRIVGHGATYDARPRTR
MYIKGAQAPVYDNRVTIKNTIILAQNLLPAFAKIEPIETMDIQVPACKGDNC
PIDAQRSLPEQDKSIVIARDVYRNLFRTVLANEGTERKRRFGQIFMGKGLMI
PKHFLYMEHYETFGSENGFVLENSNLSPLIPCSVLQGNHVDPRDIAVQ
LPNIGFQAQAYKHIIQDOPLARVSDPAILARVONASRRRQGTNYREIIPMLSTA
TPEDHUYESTVPEANVYVONRGSTYAAVYFPGCGSILIASNAIATOKIMGACI
THMKQISVLRQIIRLMKROPISQYGEHIEPLVDNPDILKENGTFLYGBELR
RINGSVYTALEKSPAYGKLIESPKPYLPLPDENSTTIDPMLQNRKGVVYVS
QQRVADAYEAMSVFYQREYINSPAHVYLPTRRETIIVGIDNPPININRTOAGAPY
TYEKEGSGKMGKMGMDYDLINDACKQALMDEIAGQCLDINVRRIIYIDLKDA
KIPLAKANVKETRLFTACPLHYTLIPQYVFIYFAHAMRNRVQNSIIVAGNMBEMD
LAKRLAKRNKHVIAQVSNPDGTLPOVYEVNAVITVDMPEMRWDDIIVADRIVNG
RELDYDEFOPLMKIGVGCINHLIHAHSDITGAALLYYRANGIPSGCPATILNSIV
NHCCLASMDINDGTPYATMNSFEFTSSIFYGDGPIINIRPEVIDIYNOETLTPVL
KKNLEMTMDIAGKCECVKARTEVSPFLKRPFSFVGLWAPIDIVLIDAPNVY
RIGNQLEPLRICVDLTISGLTELMHNDKVTDAKMTKMINGLDTRGTGLEFNDSRS
TTLLKLNBERLGGDFEINY"

<6820..9243
/note="encodes CP1, CP2 and CP3; start codon not
determined"

/codon_start=1
/product="capsid protein precursor"
/protein_id="AA01471.1"
/db_xref="GI:24431320"
/translation="PAAVFOETQMTHEQQQLITTSSEGMTSTSYTDPLDMSYLS
VDDGRNHSIIDFLQRPINQNIEMSTNDNAGKTLMAVDLPDLTINNSAKCEKRY
GFADVDELKQVNAOPQOAGRLILVLPKXYIGEDQKQYDDRTINNSAKCEKRY
PRYDLIDSTCTEAMTCVYSPYLSFSLTNGVYIGFQKVVVSPVLVDSGSLVLC
LMINFKIKIKIPAMPATIAATAGVTAIDSGAGIYSAASVSVALSPLOVPYV
SYARPAKWSNTIRIDVAKHFGMSKPTTAAAPHLNLTGSRFMANADVGMSTSLG
AISLEBNPLTRTIDEMTVAAIARPCIKRQTSNGKAGDVLYATPSTSPST
AVSTRAVAPSHLAVIISPFMTMRGGINFRKFPVTKFSGRVRLFPVGDVSNQKLP
DNADPNASVSVVDLRSDITVTRVNPVSQVPMKLSADVTSPKDOHSGRLVYL
LNELRATSVSDTISCLVEVSGADSPFLSPQPKIYPTLRAKSTLTNVRIGIADV
NVASFPVSPEIIOKGEVSGSKROQGSFTSALTVEGRTLSLOILKRFILYS
NLTSTKNNNLNRYINSEKTPKPIAAQQTQNTVNIIDYVSYVIFAYRGSPPKLA
KVTAAKIRLIPENGATEANAGVQEDNTVVIDETRTADVMPRLBETFEQVPHYSR
YPLPNTAGSIPIYVQDILLQRLNVSVNLTEATTKALFYRAVGDDPSFCGLGPP
VTRCTNPLK"

ORIGIN

Query Match 19.8%; Score 36.8; DB 13; Length 9812;
Best Local Similarity 37.5%; Pred. No. 2.9;
Matches 42; Conservative 23; Mismatches 47; Indels 0; Gaps 0;

Qy 19 UUGCUAUAUAAAUUCUGUACUAAAGUUAUUGCUAUAUAGUUGCGCUG 78

Dd		6656 TTAGGTTAATTAAGCTTATGAACATAAAAAATTACTCTTAATGCATTGTATGTTAAAGCCTTA	6715
Oy		79 CCUAAUUUGACAUCUUCUCAGAGUAGCGCCGUUGCAAGUCACCAAGAUCGAG	130
Dd		6716 GTTATTTAACTTACTTATTCAMAGATGGCGGTGGAGGCCCAAGAACTCTAG	6767
RESULT 5			
LOCUS	AX763949	9829 bp	RNA linear PAT 25-JUN-2003
DEFINITION	Sequence 1 from Patent WO03040177.		
VERSION	AX763949		
KEYWORDS	AX763949.1 GI:32258304		
SOURCE	Aphid lethal paralysis virus Aphid lethal paralysis virus Viruses; ssRNA positive-strand viruses, no DNA stage; Dicistroviridae; Crispavirinae.		
REFERENCE	van der Wilk,F. and van Munster,M. Nucleic acid derived from an ALPV virus, compositions containing said nucleic acid and uses thereof Patent: WO 03040177-A 1 15-MAY-2003;		
JOURNAL	Plant Research International B.V. (NL) Location/Qualifiers 1. .9829		
FEATURES	/organism="Aphid lethal paralysis virus" /mol_type="unassigned RNA" /db_xref="taxon:209529"		
ORIGIN			
Query Match	19.8%; Score 36.8; DB 6; Length 9829;		
Beef Local Similarity	37.5%; Pred. No. 2.9;		
Matches	42; Conservative 23; Mismatches 47; Indels 0; Gaps 0;		
Oy		19 UUUCCGAUAUAAAUCUGUCAUAAAAGCGCAAAGUAGUAGCUAUGAGUAGCGCGU	78
Dd		6656 TTAGGTTAATTAAGCTTATGAACATAAAAAATTGCTTATGCTTATGTTAAGCCTTA	6715
Oy		79 CCUAAUUUGACAUCUUCUCAGAGUAGCGCCGUUGCAAGUCACCAAGAUCGAG	130
Dd		6716 GTTATTTAACTTACTTATTCAMAGATGGCGGTGGAGGCCCAAGAACTCTAG	6767
RESULT 6			
LOCUS	BD173513	189 bp	RNA linear PAT 18-FEB-2003
DEFINITION	Novel tertiary structure having ability to accelerate translation activity.		
ACCESSION	BD173513		
VERSION	BD173513.1 GI:20414844		
KEYWORDS	WO 02061080-A/3. Drosophila C virus Drosophila C virus Viruses; ssRNA positive-strand viruses, no DNA stage; Dicistroviridae; Crispavirinae.		
SOURCE	Nakashima,N. and Kanamori,Y. 1 (bases 1 to 189)		
ORGANISM	Novel tertiary structure having ability to accelerate translation activity Patent: WO 02061080-A 3 08-AUG-2002;		
REFERENCE	JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SCIENTIFIC AND TECHNOLOGICAL SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA,YASUSHI KANAMORI		
AUTHORS	OS Drosophila C virus		
TITLE	PD WO 02061080-A/3		
JOURNAL	PN 08-AUG-2002		
COMMENT	PR 31-JAN-2001 WO 2001JP000641		
	PI 25-JAN-2001 JP 01P 016746		
	PC NOBUHIKO NAKASHIMA,YASUSHI KANAMORI		
	PG C12N15/11,C12N15/86,C12P21/02		
	CC Novel tertiary structure having ability to accelerate CC		

[illegible]

Qy	60	AUAGUAAAGUGCGCUCUCCUUAUAGCAUUCUCUCAGAGUGCGCGUGCAGUCCA	119
Db	62	TATCTTAATAATTAAGTTAACTATTAGTTTAACTGTTCAAGATCCTATTGGCAGCCCC	121
Qy	120	ACAAGAUCCAGG 131	
Db	122	ATAATATCCAGG 133	
RESULT 8			
LOCUS	AR496112	1345 bp	DNA
DEFINITION	Sequence 1072 from patent US 6703491.	linear	PAT 22-SEP-2004
ACCESSION	AR496112		
VERSION	AR496112.1	GI:524431587	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1345) Homburger,S.A., Ebens,A.J. Jr., Erickson,C.S., Francis-Lang,H.L., Margolis,J.S., Reddy,B.P., Ruddy,D.A. and Buchman,A.R.		
TITLE	Drosophila sequences		
JOURNAL	Patent: US 6703491-A 1072 09-MAR-2004; Exelixis, Inc.; South San Francisco, CA		
FEATURES	Location/Qualifiers		
source	1..1345 /organism="unknown" /mol_type="genomic DNA"		
ORIGIN			
Query Match	19.6%; Score 36.4; DB 6; Length 1345;		
Best Local Similarity	38.6%; Pred. No. 4.5;		
Matches	51; Conservative 28; Mismatches 51; Indels 2; Gaps 1;		
Qy	2	UGACUAGUGAUCUCUCUUCGUAAUAAAUUCUGUACAUAAAAGUGCAAGUAGUUCU--	59
Db	518	TTAAGATGTGATCTTGCTCTCTTATACATTTTGAGAGTTATAGAAAGAGTAGTGC	577
Qy	60	AUAGUAAAGUGCGCUCUCCUUAUAGCAUUCUCUCAGAGUGCGCGUGCAGUCCA	119
Db	578	TATTTTAATAATTAAGTTAACTATTAGTTTAACTGTTCAAGATCCTATTGGCAGCCCC	637
Qy	120	ACAAGAUCCAGG 131	
Db	638	ATAATATCCAGG 649	
RESULT 9			
LOCUS	AR511394	1345 bp	DNA
DEFINITION	Sequence 16354 from patent US 6703491.	linear	PAT 22-SEP-2004
ACCESSION	AR511394		
VERSION	AR511394.1	GI:52446869	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1345) Homburger,S.A., Ebens,A.J. Jr., Erickson,C.S., Francis-Lang,H.L., Margolis,J.S., Reddy,B.P., Ruddy,D.A. and Buchman,A.R.		
TITLE	Drosophila sequences		
JOURNAL	Patent: US 6703491-A 16354 09-MAR-2004; Exelixis, Inc.; South San Francisco, CA		
FEATURES	Location/Qualifiers		
source	1..1345 /organism="unknown" /mol_type="genomic DNA"		
ORIGIN			
Query Match	19.6%; Score 36.4; DB 6; Length 1345;		
Best Local Similarity	38.6%; Pred. No. 4.5;		
Matches	51; Conservative 28; Mismatches 51; Indels 2; Gaps 1;		

[illegible]

Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA

and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KZIO
Center clone name: CH230-unknown
----- Summary Statistics
Assembly program: Phrap, version 0.990329
Consensus quality: 225198 bases at least Q40
Consensus quality: 227434 bases at least Q30
Consensus quality: 228930 bases at least Q20
Estimated insert size: 230911; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 58717: contig of 58717 bp in length
* 58718 58817: gap of unknown length
* 58818 239102: contig of 180285 bp in length.
Location/Qualifiers

1. 239102
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-unknown"
1. 1145
/note="wgs contig"
58718. 58817
/estimated_length=unknown
58818. 60157
/note="wgs contig"
91380. 93878
/note="wgs contig"

ORIGIN

Query Match 18.8%; Score 35; DB 14; Length 239102;
Best Local Similarity 36.4%; Pred. No. 7.9;
Matches 36; Conservative 23; Mismatches 40; Indels 0; Gaps 0;

Oy 1 UGACUUVGUGAUCUGCUAUAUUAUUCUGAUAUAUUGCAUAUUGCUA 60
Db 47807 TTATCTTTGGCTTTCTTATCTTAAATTCACCTTCAGAACCAAGTATGCTCA 47748

Oy 61 UAGUUAAGUGUGCGCUUCGUUUUUAUGCAUAUUCUCA 99
Db 47747 TAGTGAATCATCTCTGGGCATACATATCTCTCTCA 47709

RESULT 14
AC137461/c 267809 bp DNA linear HTG 20-NOV-2002
LOCUS AC137461
DEFINITION Rattus norvegicus clone CH230-unknown, ** SEQUENCING IN PROGRESS

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISMREFERENCE
AUTHORS

***, 3 unordered pieces.
AC137461
AC137461.1 GI:25138600
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
Rattus norvegicus
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Rattus.
1 (bases 1 to 267809)

Munzy,D,Marie, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angiano,D, Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Biwalo,K, Blair,J, Blankenburg,K, Blych,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,B, Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Cente,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Kocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gebregeorgis,E, Geer,K, Gill,R, Gaddy,M, Guerra,W, Guevara,M, Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M, Hollins,B, Howells,S, Huliyk,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowic,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorenshuwa,L, Louisedge,H, Lozada,R,J, Lu,X, Ma,J, Maneshwari,M, Mahindartine,M, Mahmood,M, Malloy,K, Mangum,A, Mangun,B, Mapa,P, Martin,K, Martin,R, Martinez,B, Mawhney,S, McLeod,M,P, McNeill,T,Z, Meenen,E, Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwakoilemeh,O, Okwomu,G, Olarnpunsagoon,A, Pal,S, Parks,K, Paeternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C, Plopper,F, Polndexter,A, Popovic,D, Primus,E, Pu,L,L, Puzo,M, Quiroz,J, Rachin,E, Reeves,K, Regier,M,A, Reigh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rivers,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savary,G, Scherer,S, Scott,G, Shatman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Sleson,I, Sitter,C,D, Smales,D, Sneed,A, Sodergren,E, Song,X-Z, Sorelle,R, Sosa,J, Steinle,M, Strong,R, Sutton,A, Syatek,A, Taber,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Umani,K, Valas,R, Vera,V, Villaseana,D, Waldron,L, Walker,B, Wang,J, Wang,O, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willson,R, Wleczyk,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Niederhausern,A, Weise,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstein,G, and Gibbs,R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished
2 (bases 1 to 267809)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KZIN
Center clone name: CH230-unknown
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 163420 bases at least Q40
Consensus quality: 168676 bases at least Q30
Consensus quality: 172110 bases at least Q20
Estimated insert size: 157250; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 265269: contig of 265269 bp in length
* 265270 265369: gap of unknown length
* 265370 266601: contig of 1232 bp in length
* 266602 266701: gap of unknown length
* 266702 267809: contig of 1108 bp in length.
Location/Qualifiers

FEATURES

source 1..267809
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-unknown"
1..2027
misc_feature /note="wgs contig"
16011..17501
misc_feature /note="wgs contig"
20015..21050
misc_feature /note="wgs contig"
51039..52759
misc_feature /note="wgs contig"
90217..91472
misc_feature /note="wgs contig"
95100..96591
misc_feature /note="wgs contig"
108632..110236
misc_feature /note="wgs contig"
123509..124765
misc_feature /note="wgs contig"
130438..131667
misc_feature /note="wgs contig"
142596..143821
misc_feature /note="wgs contig"
146719..147808
misc_feature /note="wgs contig"
180964..182519
misc_feature /note="wgs contig"
211046..212894
misc_feature /note="wgs contig"
246154..247283
misc_feature /note="wgs contig"
265270..265369
gap /estimated_length=unknown

gap

266602..266701
/estimated_length=unknown

ORIGIN

Query Match

Best local similarity 36.4%; Pred. No. 7.8; Length 267809;

Matches 36; Conservative 23; Mismatches 40; Indels 0; Gaps 0;

QY

1 UGACUAGUAGUAGUCUUCUUGUAAUUAUUCUGUACAUAAAAGUGAAUUGUCUA 60
45101 TTACTTGTGGCTGTTCTTCAATTAATTAATTCATTCATTCACAAAGCAAAAGATGTCTA 45042

Db

61 UAGUUAAGUUGCGUCCUUAUUUAGGCAUACUUCUA 99
45041 TAGTGAATCATTTCTGGCAATACATAATCTCTCA 45003

RESULT 15

LOCUS

BX784037 126572 bp DNA linear VRT 17-SEP-2004
Zebrafish DNA sequence from clone DKEY-96H14 in linkage group 22,
complete sequence.

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: zf1sh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest, except on the rare

occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep

zebrafish pUC subclones occasionally display inconsistency over the

length of mononucleotide A/T runs and conserved TA repeats. Where

this is found the longest good quality representation will be

submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat

discovery system (Zhirong Bao and Sean Eddy, submitted), and those

beginning 'dir' were identified by Rick Waterman (Stephen Johnson

lab, Maebui). For further information see

http://www.sanger.ac.uk/Projects/D_reio/fishmask.shtml DKEY-96H14

is from a Zebrafish BAC library

VECTOR: pindigobAC-5.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 04:43:02 ; Search time 268.794 Seconds
(without alignments)
4611.840 Million cell updates/sec

Title: US-10-088-750C-5

Perfect score: 186
Sequence: 1 uugcaucuaugauucuuugcu.....uugcaucuaagguugacuc 186

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N_Geneseq_21:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003as:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004as:*
- 13: Geneseq2004bs:*
- 14: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	186	100.0	186	AB280711	AB280711 Triatoma
2	186	100.0	198	AA150548	AA150548 Triatoma
3	36.8	19.8	9829	AA160324	AA160324 Aphid let
4	36.4	19.6	189	AB280709	AB280709 Drosophila
5	36.4	19.6	201	AA150546	AA150546 Drosophila
6	33.6	18.1	188	AB280710	AB280710 Cricket p
7	33.6	18.1	200	AA150547	AA150547 Cricket p
8	33.6	18.1	854	AB033052	AB033052 Oligonuc
9	33.6	18.1	854	AB033053	AB033053 Oligonuc
10	33.2	17.8	419	ABT22167	ABT22167 Breast ca
11	31	16.7	5278	AA150548	AA150548 Human imm
12	31	16.7	5278	AA150548	AA150548 Human imm
13	30.6	16.5	874	AA208975	AA208975 Mouse gen
14	30.6	16.5	38977	AA208975	AA208975 Mouse gen
15	30.4	16.3	667	ADK53988	ADK53988 Plant DNA
16	30.2	16.2	2960	AA141616	AA141616 Adzuki be
17	30	16.1	478	AA138538	AA138538 Human toe
18	30	16.1	478	AA138538	AA138538 Probe #72
19	30	16.1	478	AAK32723	AAK32723 Human bon

20	30	16.1	478	AAK06986	AAK06986 Human bra
21	30	16.1	478	AB280711	AB280711 Human liv
22	30	16.1	478	AB280711	AB280711 Human gen
23	30	16.1	492	ABV46608	ABV46608 Human pro
24	29.8	16.0	941	AA124946	AA124946 Human bre
25	29.8	16.0	1296	ABK75374	ABK75374 Bacillus
26	29.6	15.9	187	AB280708	AB280708 Himetobi
27	29.6	15.9	199	AA150545	AA150545 Himetobi
28	29.6	15.9	482	ACH44858	ACH44858 Human foe
29	29.6	15.9	556	AAH11148	AAH11148 Human cdn
30	29.6	15.9	1301	ACA03978	ACA03978 CDNA down
31	29.6	15.9	1558	AAQ10896	AAQ10896 Encodes X
32	29.6	15.9	2290	AA29606	AA29606 Human end
33	29.6	15.9	3022	AAH17724	AAH17724 Human cdn
34	29.6	15.9	3340	ABL18096	ABL18096 Drosophila
35	29.6	15.9	5846	AAAD49503	AAAD49503 Human TRI
36	29.6	15.9	5966	ADL33372	ADL33372 Human tria
37	29.6	15.9	6718	AB557422	AB557422 Human pro
38	29.6	15.9	46718	ADG62974	ADG62974 Human pro
39	29.6	15.9	110000	AD246976_1	AD246976 (2 of
40	29.6	15.9	24987	ABN85733	ABN85733 Mouse gen
41	29.4	15.8	430	AAV87077	AAV87077 EST clone
42	29.4	15.8	688	ADK65104	ADK65104 Plant ful
43	29.4	15.8	1164	ADA70696	ADA70696 Rice gene
44	29.4	15.8	1746	ADJ39578	ADJ39578 Plant cdn
45	29.4	15.8	73334	ABL34125	ABL34125 Human imm

ALIGNMENTS

RESULT 1
ID AB280711 standard; RNA, 186 BP.
XX AB280711;
XX AC
XX DT 15-OCT-2003 (first entry)
XX DE Triatoma virus derived pseudoknot sequence.
XX KW Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;
XX albumen; impurity; higher-order structure; intergenic region; IGR-IRIS;
XX internal ribosome entry site; ss.
XX OS
XX PN WO2003033719-A1.
XX PD 24-APR-2003.
XX PF 08-OCT-2002; 2002WO-JP010447.
XX PR 17-OCT-2001; 2001JP-00319923.
XX PA (NAG-) NAT INST AGRONOMIC SCI.
XX (NAKE-) WAKENYAKU KK.
XX Nakashima N, Shibuya N, Nishikawa S;
XX WPI; 2003-403230/38.
XX Cell-free protein synthesis means in wheatgerm system to establish
XX overexpression of target gene with base sequence sustaining translation
XX activity and function promotion, for producing useful proteins.
XX Claim 1; Page 33; 39pp; Japanese.
XX The invention relates to a cell-free protein synthesis system derived
XX from wheatgerm where there is substantial exclusion of wheatgerm embryo
XX albumen impurities. The novel system uses a sequence having a higher-
XX order RNA structure that promotes translation activity. The higher-order
XX sequence is preferably a "pseudoknot", especially derived from a range of

CC viruses (AB280707-AB280713). This sequence represents the "pseudoknot"
 CC higher-order sequence from the Triatoma virus. The sequence is used in a
 CC construct which may also include an intergenic region and internal
 CC ribosome entry site (IGR-IRBS). The method is applicable in producing
 CC useful proteins
 CC
 SQ Sequence 186 BP; 49 A; 35 C; 42 G; 0 T; 60 U; 0 Other;

Query Match 100.0%; Score 186; DB 8; Length 186;
 Best Local Similarity 100.0%; Pred. No. 1,1e-54;
 Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 UGACUAGUGAUCUUCGUAUAAAUCUGACAUAAAAGCGAAAGUAUUGCUA 60
 |||||
 Db 1 UGACUAGUGAUCUUCGUAUAAAUCUGACAUAAAAGCGAAAGUAUUGCUA 60
 |||||
 QY 61 UAGUUAAGGUGCGCGUCCUUAUUAGCGAUAUCUUCGAGUAGCGGUGCAGUCCAA 120
 |||||
 Db 61 UAGUUAAGGUGCGCGUCCUUAUUAGCGAUAUCUUCGAGUAGCGGUGCAGUCCAA 120
 |||||
 QY 121 CAAGAUCCAGGACUGUACAGAAUUUUCUUAACUCCAGUGCGGUAUUGAUAUAGU 180
 |||||
 Db 121 CAAGAUCCAGGACUGUACAGAAUUUUCUUAACUCCAGUGCGGUAUUGAUAUAGU 180
 |||||
 QY 181 UGACUC 186
 |||||
 Db 181 UGACUC 186

RESULT 2

AA150548
 ID AAL50548 standard; RNA; 198 BP.

AC AAL50548;
 XX 19-DEC-2002 (first entry)
 DT Triatoma virus RNA sequence.
 DE Triatoma virus RNA sequence.
 XX CrPV-like virus; ss, higher-order structure; drug development;
 KM drug production; translational activity-promoting function;
 KW protein synthesis; structural analysis.

OS Triatoma virus.
 XX WO200261080-A1.
 PN 08-AUG-2002.
 PD 31-JAN-2001; 2001WO-JP000641.
 PF 25-JAN-2001; 2001JP-00016746.
 PR (NAG-) NAT INST AGROBIOLOGICAL SCI.
 PA Nakaishima N, Kanamori Y;
 PI WPI; 2002-627482/67.
 DR

XX Translational activity-promoting higher-order structure of CrPV-like
 PT viruses for protein translation when suitably initiated, useful in
 PT synthesis of proteins and polypeptides of foreign species for application
 PT in drugs.

XX Claim 1; Fig 1-2; 38pp; Japanese.

XX The invention comprises seven RNA sequences (CrPV-like viruses) which
 CC have a higher-order structure that sustains translational activity-
 CC promoting function. The RNA sequences of the invention are useful in the
 CC synthesis of proteins and polypeptides for application in developing and
 CC producing drugs. The RNA sequences of the invention are also useful in
 CC basic research of protein synthesis and structural analysis by the gene
 CC recombinant technique. The present nucleotide represents a Triatoma virus

CC RNA sequence of the invention

XX Sequence 198 BP; 54 A; 36 C; 44 G; 0 T; 64 U; 0 Other;

SQ Query Match 100.0%; Score 186; DB 6; Length 198;
 Best Local Similarity 100.0%; Pred. No. 1,1e-54;
 Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 UGACUAGUGAUCUUCGUAUAAAUCUGACAUAAAAGUGCAAGUAUUGCUA 60
 |||||
 Db 1 UGACUAGUGAUCUUCGUAUAAAUCUGACAUAAAAGUGCAAGUAUUGCUA 60
 |||||
 QY 61 UAGUUAAGGUGCGCGUCCUUAUUAGCGAUAUCUUCGAGUAGCGGUGCAGUCCAA 120
 |||||
 Db 61 UAGUUAAGGUGCGCGUCCUUAUUAGCGAUAUCUUCGAGUAGCGGUGCAGUCCAA 120
 |||||
 QY 121 CAAGAUCCAGGACUGUACAGAAUUUUCUUAACUCCAGUGCGGUAUUGAUAUAGU 180
 |||||
 Db 121 CAAGAUCCAGGACUGUACAGAAUUUUCUUAACUCCAGUGCGGUAUUGAUAUAGU 180
 |||||
 QY 181 UGACUC 186
 |||||
 Db 181 UGACUC 186

RESULT 3

AA160324
 ID AAL60324 standard; RNA; 9829 BP.

AC AAL60324;
 XX 27-AUG-2003 (first entry)
 DT Aphid lethal paralysis virus (ALPV) RNA.
 DE Aphid lethal paralysis virus (ALPV) RNA.
 XX ALPV; Aphid lethal paralysis virus; homopterous insect; floriculture;
 KM aphid; white fly; agriculture; gene therapy; ss.

OS Aphid lethal paralysis virus.
 XX WO2003040177-A2.
 PN 15-MAY-2003.
 PD 25-OCT-2002; 2002WO-NL000682.
 PF 25-OCT-2001; 2001NL-01019225.
 PR (PLAN-) PLANT RES INT BV.
 PA Van Der Wilk F, Van Munster M;
 PI WPI; 2003-441523/41.
 DR

XX Novel nucleic acid derived from aphid lethal paralysis virus, useful for
 PT detecting ALPV-viruses, for preparing a monoclonal antibody against
 PT antigen, in research concerning aphids and for controlling aphids.
 PT
 PS Claim 1; Page 10-13; 17pp; English.

XX The invention relates to a novel RNA derived from an ALPV (Aphid lethal
 CC paralysis virus). ALPV RNA is useful for controlling homopterous insects
 CC such as aphids or white flies. It is useful in the research concerning
 CC aphids and for controlling aphids and in the field of agriculture or
 CC floriculture. It is also useful for detecting ALPV-viruses using PCR and
 CC for preparing a monoclonal antibody against antigen which is synthesised
 CC in vitro. The invention is also useful in gene therapy. The present
 CC sequence is ALPV RNA

SQ Sequence 9829 BP; 3083 A; 1906 C; 1879 G; 0 T; 2961 U; 0 Other;

Query Match 19.8%; Score 36.8; DB 8; Length 9829;
 Best Local Similarity 58.0%; Pred. No. 0.093;

Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 19 UUTCGUAAUAAACUCUGACUAAAAAGUCGAAAGUATUGCUAAUGUAGGUCGCGUG 78
DB 6656 UUAGGUUAUUAUGUGAGACUAAUAAAAUUAUGCUAAUGCAUUAAGUAAUUAAGGCUA 6715

QY 79 CCUUAUUAGGCAUACUCUCUGAGUUGGCGUGGCGAGUCGACAGUCCAAAGAUCCAG 130
DB 6716 GUUAUUUAACUUAUCUUAUACAAGAGGCGUGGCGAGCCCAAGAAUUCUG 6767

RESULT 4
AB280709 ID AB280709 standard; RNA; 189 BP.
XX AC AB280709;
XX DT 15-OCT-2003 (first entry)
XX DE Drosophila C virus derived pseudoknot sequence.
XX KM Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;
XX KM albumen; impurity; higher-order structure; intergenic region; IGR-IRBS;
XX KM internal ribosome entry site; ss.
OS Drosophila C virus.
XX PN WC2003033719-A1.
XX PD 24-APR-2003.
XX PF 08-OCT-2002; 2002WC-JP010447.
XX PR 17-OCT-2001; 2001JP-00319923.
XX PA (NAG-) NAT INST AGROBIOLOGICAL SCI.
XX PA (WAKE-) WAKENYAKU KK.
XX PI Nakashima N, Shibuya N, Nishikawa S;
XX DR WPI; 2003-403230/38.
XX CC Cell-free protein synthesis means in wheatgerm system to establish
PT overexpression of target gene with base sequence sustaining translation
PT activity and function promotion, for producing useful proteins.
PS Claim 1; Page 32; 39pp; Japanese.
XX CC The invention relates to a cell-free protein synthesis system derived
CC from wheatgerm where there is substantial exclusion of wheatgerm embryo
CC albumen impurities. The novel system uses a sequence having a higher-
CC order RNA structure that promotes translation activity. The higher-order
CC sequence is preferably a "pseudoknot", especially derived from a range of
CC viruses (AB280707-AB280713). This sequence represents the "pseudoknot"
CC higher-order sequence from the Drosophila C virus. The sequence is used
CC in a construct which may also include an intergenic region and internal
CC ribosome entry site (IGR-IRBS). The method is applicable in producing
CC useful proteins
XX

SQ Sequence 189 BP; 57 A; 30 C; 34 G; 0 T; 68 U; 0 Other;

Query Match 19.6%; Score 36.4; DB 8; Length 189;
Best Local Similarity 59.8%; Pred. No. 0.027; Mismatches 51; Indels 2; Gaps 1;
Matches 79; Conservative 0; Mismatches 51; Indels 2; Gaps 1;

QY 2 UGACUAGUGUAGUCUGUCUUUGUAAUAAUCUGUACUAAAAAGUCGAAAGUATUGCU-- 59
DB 2 UUAAGAGUGUAGUCUGUCUUUGUAAUAAUCUGUAGAGUUAUAAAGAGAAUAGUAGUC 61

QY 60 AUAGUUAAGUGGCGUGGCUUAUUUAGCAUACUUCUCAGAGUUGGCGGUGGCAUCCA 119
DB 62 UAUCUUUAUUAUUAUGGUUAACUUAUUUAGUUUACUGUUCAGAGUUGCUUAGGAGCCCC 121

QY 120 ACAAGUCCAGG 131
DB 122 AUAUAUCCAGG 133

RESULT 5
AAL50546 ID AAL50546 standard; RNA; 201 BP.
XX AC AAL50546;
XX DT 19-DEC-2002 (first entry)
XX DE Drosophila C virus RNA sequence.
XX KM CrPV-like virus; ss; higher-order structure; drug development;
XX KM drug production; translational activity-promoting function;
XX KM protein synthesis; structural analysis.
XX OS Drosophila C virus.
XX EN WC200261080-A1.
XX PD 08-AUG-2002.
XX PF 31-JAN-2001; 2001WC-JP000641.
XX PR 25-JAN-2001; 2001JP-00016746.
XX PA (NAG-) NAT INST AGROBIOLOGICAL SCI.
XX PI Nakashima N, Kanamori Y;
XX DR WPI; 2002-627482/67.
XX CC Translational activity-promoting higher-order structure of CrPV-like
PT viruses for protein translation when suitably initiated, useful in
PT synthesis of proteins and polypeptides of foreign species for application
PT in drugs.
XX PS Claim 1; Fig 1-2; 38pp; Japanese.
XX CC The invention comprises seven RNA sequences (CrPV-like viruses) which
CC have a higher-order structure that sustains translational activity-
CC promoting function. The RNA sequences of the invention are useful in the
CC synthesis of proteins and polypeptides for application in developing and
CC producing drugs. The RNA sequences of the invention are also useful in
CC basic research of protein synthesis and structural analysis by the gene
CC recombinant technique. The present nucleotide represents a Drosophila C
CC virus RNA sequence of the invention
XX

SQ Sequence 201 BP; 61 A; 33 C; 35 G; 0 T; 72 U; 0 Other;

Query Match 19.6%; Score 36.4; DB 6; Length 201;
Best Local Similarity 59.8%; Pred. No. 0.028; Mismatches 51; Indels 2; Gaps 1;
Matches 79; Conservative 0; Mismatches 51; Indels 2; Gaps 1;

QY 2 UGACUAGUGUAGUCUGUCUUUGUAAUAAUCUGUACUAAAAAGUCGAAAGUATUGCU-- 59
DB 2 UUAAGAGUGUAGUCUGUCUUUGUAAUAAUCUGUAGAGUUAUAAAGAGAAUAGUAGUC 61

QY 60 AUAGUUAAGUGGCGUGGCUUAUUUAGCAUACUUCUCAGAGUUGGCGGUGGCAUCCA 119
DB 62 UAUCUUUAUUAUUAUGGUUAACUUAUUUAGUUUACUGUUCAGAGUUGCUUAGGAGCCCC 121

QY 120 ACAAGUCCAGG 131
DB 122 AUAUAUCCAGG 133

RESULT 6
AB280710 ID AB280710 standard; RNA; 188 BP.


```
XX AC ABZ80710;
XX XX
XX DT 15-OCT-2003 (first entry)
XX DE Cricket paralysis virus derived pseudoknot sequence.
XX XX
XX KM Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;
XX KM albumen; impurity; higher-order structure; intergenic region; IGR-IRRS;
XX KM internal ribosome entry site; ss.
XX XX
XX OS Cricket paralysis virus.
XX PN MO2003033719-A1.
XX PD 24-APR-2003.
XX PF 08-OCT-2002; 2002WO-JP010447.
XX PR 17-OCT-2001; 2001JP-00319923.
XX XX
XX PA (NAG-) NAT INST AGRBIOLOGICAL SCT.
XX PA (WAKE-) WAKENTAKU KK.
XX PI Nakashima N, Shibuya N, Nishikawa S;
XX DR WPI; 2003-403230/38.
XX XX
XX PT Cell-free protein synthesis means in wheatgerm system to establish
XX PT overexpression of target gene with base sequence sustaining translation
XX PT activity and function promotion, for producing useful proteins.
XX PS
XX PS Claim 1; Page 33; 39pp; Japanese.
XX XX
XX CC The invention relates to a cell-free protein synthesis system derived
XX CC from wheatgerm where there is substantial exclusion of wheatgerm embryo
XX CC albumen impurities. The novel system uses a sequence having a higher-
XX CC order RNA structure that promotes translation activity. The higher-order
XX CC sequence is preferably a "pseudoknot", especially derived from a range of
XX CC viruses (ABZ80707-ABZ80713). This sequence represents the "pseudoknot"
XX CC higher-order sequence from the Cricket paralysis virus. The sequence is
XX CC used in a construct which may also include an intergenic region and
XX CC internal ribosome entry site (IGR-IRRS). The method is applicable in
XX CC producing useful proteins
XX SQ
XX Sequence 188 BP; 56 A; 33 C; 36 G; 0 T; 63 U; 0 Other;
XX
XX Query Match 18.1%; Score 33.6; DB 8; Length 188;
XX Best Local Similarity 70.3%; Pred. No. 0.26;
XX Matches 45; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX
XX QY 70 UUGCGCUCUUGCCUUAUUGGCAUACUCUCAGAGUGGCGGUGGCAUACAGAUCCA 129
XX XX
XX DB 71 UUAAGGUAUGCUAUUUAUGCUUAUGCUUACGUCACGAGUCCUUAUGGACCCCAUAUCCA 130
XX
XX QY 130 GGGA 133
XX DB 131 GGAA 134
XX
XX RESULT 7
XX ID AAL50547 standard; RNA; 200 BP.
XX AC AAL50547;
XX XX
XX DT 19-DEC-2002 (first entry)
XX DE Cricket paralysis virus RNA sequence.
XX XX
XX KM CrPV-like virus; ss; higher-order structure; drug development;
XX KM drug production; translational activity-promoting function;
XX KM protein synthesis; structural analysis.
```

```
XX OS Cricket paralysis virus.
XX XX
XX PN MO200261080-A1.
XX XX
XX PD 08-AUG-2002.
XX XX
XX PF 31-JAN-2001; 2001WO-JP000641.
XX PR 25-JAN-2001; 2001JP-00016746.
XX XX
XX PA (NAG-) NAT INST AGRBIOLOGICAL SCT.
XX PI Nakashima N, Kanamori Y;
XX DR WPI; 2002-627482/67.
XX XX
XX PT Translational activity-promoting higher-order structure of CrPV-like
XX PT viruses for protein translation when suitably initiated, useful in
XX PT synthesis of proteins and polypeptides of foreign species for application
XX PT in drugs.
XX PS
XX PS Claim 1; Fig 1-2; 38pp; Japanese.
XX XX
XX CC The invention comprises seven RNA sequences (CrPV-like viruses) which
XX CC have a higher-order structure that sustains translational activity-
XX CC promoting function. The RNA sequences of the invention are useful in the
XX CC synthesis of proteins and polypeptides for application in developing and
XX CC producing drugs. The RNA sequences of the invention are also useful in
XX CC basic research of protein synthesis and structural analysis by the gene
XX CC recombinant technique. The present nucleotide represents a Cricket
XX CC paralysis virus RNA sequence of the invention
XX SQ
XX Sequence 200 BP; 60 A; 36 C; 37 G; 0 T; 67 U; 0 Other;
XX
XX Query Match 18.1%; Score 33.6; DB 6; Length 200;
XX Best Local Similarity 70.3%; Pred. No. 0.26;
XX Matches 45; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX
XX QY 70 UUGCGCUCUUGCCUUAUUGGCAUACUCUCAGAGUGGCGGUGGCAUACAGAUCCA 129
XX XX
XX DB 71 UUAAGGUAUGCUAUUUAUGCUUAUGCUUACGUCACGAGUCCUUAUGGACCCCAUAUCCA 130
XX
XX QY 130 GGGA 133
XX DB 131 GGAA 134
XX
XX RESULT 8
XX ID ABQ33052 standard; DNA; 854 BP.
XX AC ABQ33052;
XX XX
XX DT 12-UTL-2002 (first entry)
XX DE
XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 19643.
XX XX
XX KM Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX KM drug; side effect; cancer; central nervous system; cardiovascular;
XX KM gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KM SNP; cell differentiation; ds.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO200218632-A2.
XX PD 07-MAR-2002.
XX XX
XX PF 01-SEP-2001; 2001WO-EP010074.
XX PR 01-SEP-2000; 2000DE-01043826.
XX PR 05-SEP-2000; 2000DE-01044543.
```


PR	20-APR-2001; 2001US-0285163P.	XX
PA	(MILL-) MILLENNIUM PHARM INC.	XX
PI	Lillie J, Palermo A, Wang Y, Steinmann K, Elias J, Mertens M;	XX
XX	WPI: 2003-093053/08.	XX
DR		XX
PT	Novel isolated polypeptide encoded by breast cancer marker gene, useful	XX
PT	for diagnosing, staging, monitoring, prognosing and treating diseases	XX
PT	associated with breast cancer.	XX
PS	Disclosure; Page 163; 725pp; English.	XX
CC		XX
CC	The invention relates to an isolated polypeptide encoded by a breast	XX
CC	cancer marker gene comprising any of 1417 21-805 nucleotide sequences,	XX
CC	given in the specification. The methods of the invention are useful for	XX
CC	diagnosing patients having an identified breast mass or symptoms	XX
CC	associated with breast cancer, to diagnose breast cancer or its	XX
CC	precursors, and for monitoring the efficacy of treatment of a breast	XX
CC	cancer patient (e.g. efficacy of chemotherapy). The methods are also	XX
CC	useful for evaluating a patient before, after or during therapy, to	XX
CC	evaluate the reduction in a tumor burden. The breast cancer marker gene	XX
CC	proteins are useful as immunogens for raising antibodies, by immunising a	XX
CC	mammal with a breast cancer marker protein. The marker proteins are	XX
CC	useful as bait proteins in a two-hybrid or three-hybrid assay, to	XX
CC	identify other proteins which bind to or interact with the marker	XX
CC	proteins. The breast cancer marker genes are useful as surrogate marker	XX
CC	genes for one or more disorders, disease states or conditions leading to	XX
CC	disease states, in particular, breast cancers. The breast cancer marker	XX
CC	genes are useful as pharmacodynamic marker genes. An antibody which	XX
CC	selectively binds to a protein of a breast cancer marker gene is useful	XX
CC	for treating cancers, particularly breast cancers. The host cell of the	XX
CC	invention is useful for producing non-human transgenic animals. This	XX
CC	polynucleotide sequence represents one of the breast cancer marker genes	XX
CC	of the invention	XX
XX		XX
XX	Sequence 419 BP; 100 A; 95 C; 94 G; 126 T; 0 U; 4 Other;	XX
Query Match	17.8%; Score 33.2; DB 10; Length 419;	
Best Local Similarity	34.7%; Pred. No. 0.49;	
Matches	52; Conservative 25; Mismatches 73; Indels 0; Gaps 0	
QY	27 UAAAUCUGUACAUAAGAUGCGAAGUAGUAGUAGUAGUGCGUCGUUUA 86	
DB	145 TAAGGGTTTCACATTAATTAAACATCTATATTATTAACAACAAAGTCAATG 204	
QY	87 GGAUACUUCUUCGAGUUGCGGCGUGGACGUCCAAAGAUCCAGGACUGUACGAUUT 146	
DB	205 GGCATCTACTTGGATGTAAGTGCCTTTCTTACACCAAAATACATGTATACAAAAGATT 264	
QY	147 UCCUAAUCCUGAGUGCGGUGUGUAGAUUCA 176	
DB	265 AAACACTTTTCCGCCCGCTCTTAAAGAACTA 294	
RESULT 11		
ASAS46376		
ASAS46376	standard; DNA; 5278 BP.	
AC	ASAS46376;	
XX		
DT	18-DEC-2001 (first entry)	
XX		
DE	Tumour suppressor gene derived chemically modified sequence #98.	
XX		
KM	Human; tumour suppressor gene; oncogene; antitumour; cytosolic; cancer;	
XX	tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;	
XX	cytosine methylation; de.	
XX		
OS	Homo sapiens.	
XX		
XX	MO200168912-A2.	
XX		

XX	20-SEP-2001.
PD	
PF	15-MAR-2001; 2001WO-EP002955.
XX	
PR	15-MAR-2000; 2000DE-01013847.
PR	06-APR-2000; 2000DE-01019058.
PR	07-APR-2000; 2000DE-01019173.
PR	30-JUN-2000; 2000DE-01032529.
PR	01-SEP-2000; 2000DE-01043826.
PA	
XX	(EPIC-) EPIDENOMICS AG.
XX	
PI	Olek A, Piepenbrock C, Berlin K;
XX	
DR	WPI; 2001-602752/68.
XX	
PT	Fragments of chemically modified genes associated with tumor suppressor
PT	genes and oncogenes, useful in designing primers and probes for analyzing
PT	diseases associated with cytosine methylation state e.g. cancer.
XX	
PS	Claim 1; SEQ ID NO 98; 27pp; English.
XX	
CC	The invention relates to a nucleic acid comprising a sequence of 18
CC	bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC	bisulphite, of genes associated with tumour suppression and oncogenes
CC	having a sequence taken from 536 (actually 533 since numbers 408, 458 and
CC	500 are missing from the sequence listing) sequences (Ss) and sequences
CC	complementary to (Ss). The nucleic acid may be a peptide nucleic acid-
CC	oligomer (PNA) of at least 9 nucleotides and may form part of a set of
CC	probes for detecting the cytosine methylation state and/or single
CC	nucleotide polymorphisms and also to be used in an array for analysing
CC	diseases associated with CGG dinucleotides e.g. cancers and tumours. The
CC	probes can also be used in a method for ascertaining genetic and/or
CC	epigenetic parameters for the diagnosis and/or therapy of existing
CC	diseases or the predisposition to specific diseases, by analysing
CC	cytosine methylations. The parameters may be compared to another set of
CC	genetic and/or epigenetic parameters, the differences serving as basis
CC	for diagnosis and/or prognosis events which are disadvantageous to
CC	patients. The present sequence is one of the 533 genomic sequences
CC	derived from tumour suppressor genes and oncogenes. Sequences with even
CC	numbered Seq ID numbers are the complementary sequence of the
CC	corresponding odd numbered sequence (e.g. ID 2 and ID1, ID 536 and ID
CC	533, except for those whose partner sequence is missing). Note: The
CC	sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pat_sequences
XX	
XX	Sequence 5278 BP; 1569 A; 84 C; 1106 G; 2519 T; 0 U; 0 Other;
XX	
XX	Query Match 16.7%; Score 31; DB 4; Length 5278;
XX	Best Local Similarity 26.3%; Pred. No. 7.8;
XX	Matches 25; Conservative 30; Mismatches 40; Indels 0; Gaps 0;
QY	6 UAUUGUACUUCUGUCGUAUAAAUUCUGACUAUAAAGUAGAUUGCUAUGU 65
DB	5170 TATCTAATATATTTTTCGTAGCGAAGATTAGAGCTAATAATATGTTTTTATTTT 5229
QY	66 AAGUGUGCGUCUGUCUUNUUAAGCAUUCUUCUCG 100
DB	5230 TATGTTTTGGGTAGATATTTTATTTATTTTATG 5264
XX	
XX	RESULT 12
XX	ABL32823
ID	ABL32823 standard; DNA; 5278 BP.
AC	ABL32823;
XX	
XX	26-MAR-2002 (first entry)
DT	
XX	
XX	Human immune system associated gene SEQ ID NO: 796.

KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antihaemic; cytosolic; noctropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritis; antidiabetic; antiparasitic;
KW antitubercular; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
ds.
XX Homo sapiens.
OS
XX MO200200928-A2.
PN
XX
XX 03-JAN-2002.
PD
XX 02-JUL-2001; 2001WO-BP007537.
PE
XX 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIC-) EPIGENOMICS AG.
PA
XX Olek A. Piepenbrock C, Berlin K;
PI
XX MPI; 2002-130909/17.
DR
XX Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
XX Claim 1; SEQ ID NO 796; 32pp + Sequence Listing; German.
PS
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
CC
XX
SQ Sequence 5278 BP; 1569 A; 84 C; 1106 G; 2519 T; 0 U; 0 Other;
Query Match 16.7%; Score 31; DB 6; Length 5278;
Best Local Similarity 26.3%; Pred. No. 7.8;
Matches 25; Conservative 30; Mismatches 40; Indels 0; Gaps 0;
QY 6 VAUGUGAUCUGUCUUAUAAUUCUGUACAUAAGUGCAAGUAGUAGUAGU 65
DB 5170 TATGTAAATATTTTTCGTAGAGAAATTAAGCTAAAGATATATGTTTTTTT 5229
QY 66 AAGGUGCGCUGCCUUAUUGGCAUACUUCUAG 100
DB 5230 TATGTTTGGGTAGATATTTTATATATTTTAA 5264
RESULT 13
AA208975
ID AA208975 standard; DNA; 874 BP.
XX
AC AA208975;
XX
XX 21-OCT-1999 (first entry)
XX
XX Strawberry plant RJ39 promoter DNA.
DE
XX
KW Receptor-like; tissue-specific expression; transgenic plant; ds.
KW disease resistance; insect resistance; fruit modification; promoter.
XX
XX Fragaria vesca.
OS
XX
XX Key Location/Qualifiers
FH promoter 1..871
FT

FT /*tag= a
FT /note= "Full-length promoter for RJ39 gene"
FT misc_signal 872..874
FT /tag= b
FT /note= "ATG translational start site"
XX
XX MO9940211-A2.
XX
XX 12-AUG-1999.
PD
XX
XX 04-FEB-1999; 99WO-US002958.
XX
XX 06-FEB-1998; 98US-00020033.
XX
XX (CALJ) CALGENE LLC.
PA
XX Wilkinson JQ;
PI
XX MPI; 1999-494302/41.
DR
XX
XX DNA construct for modifying fruit bearing plant phenotype for disease
PT resistance and to increase sugar content.
PT
XX
XX Claim 5; Fig 2; 46pp; English.
PS
XX
XX This sequence represents the promoter from the RJ39 gene of strawberry
CC plants. RJ39 mRNA is expressed in receptacle tissue and expression levels
CC increase as the fruit develops to maturity. A genomic clone of the RJ39
CC coding sequence was obtained by PCR amplification of strawberry genomic
CC DNA. A 400 base radiolabelled probe was prepared from the clone and used
CC to detect the RJ39 gene in a genomic library. A clone was identified as
CC containing the promoter and the complete coding sequence. To remove the
CC portion of coding sequence, a BglII restriction site was introduced
CC upstream of the RJ39 initiation codon by PCR amplification of the
CC promoter clone with primers RJ39-3M and T7 promoter primer. Restriction
CC digestion produced a truncated RJ39 promoter able to be inserted into a
CC vector. By additional insertion of a target gene and a termination region
CC into this vector, a DNA construct can be generated. Once this construct
CC is introduced into a plant, the target gene is expressed in increasing
CC levels in receptacle tissue as the fruit matures. Using DNA constructs
CC containing this promoter, it is possible to introduce fruit-specific
CC improvements in plants where the edible fruit comprises receptacle tissue
CC e.g., strawberries, apples and pears. Such commercially and
CC agriculturally beneficial modifications may include an increase in the
CC sugar content in fruits, resistance to plant diseases and resistance to
CC insects
CC
XX
SQ Sequence 874 BP; 265 A; 142 C; 146 G; 321 T; 0 U; 0 Other;
Query Match 16.5%; Score 30.6; DB 2; Length 874;
Best Local Similarity 28.7%; Pred. No. 5.3;
Matches 45; Conservative 33; Mismatches 79; Indels 0; Gaps 0;
QY 5 CUAVUGAUCUGUCUUAUAAUUCUGUACAUAAGUGCAAGUAGUAGUAGU 64
DB 597 CGATGCATCTTTCGTGTCACATTAACATATTAACAATATCCAAATATCTATATATG 656
QY 65 UAGGUGCGCUGCCUUAUUGGCAUACUUCUAGAGGCGCGCUGCAUCCAAAG 124
DB 657 TATAATTAATATATTTTCTAAATTTTCACTTCACATGTTTAAATTTGCCAGTCAATCAT 716
QY 125 AUCCAGGACUGUACAUAUUCUUAUACUUCGAGU 161
DB 717 CTAATTACACGCTAGAGAACTCTAGTAAGAACTGGAAT 753
RESULT 14
ACN44728
ID ACN44728 standard; DNA; 38977 BP.
XX
XX ACN44728;
XX
XX 18-NOV-2004 (first entry)
XX
XX


```
XX Mouse genomic sequence mCG2727.
XX
XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX
XX Mus musculus.
XX
XX WO2003073826-A2.
XX
XX 12-SEP-2003.
XX
XX 28-FEB-2003; 2003WO-US006235.
XX
XX 01-MAR-2002; 2002US-00087192.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW;
XX
XX WPI; 2003-328604/31.
XX
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX comprises a nucleotide sequence.
XX
XX Claim 1; SEQ ID NO 1321; 0pp; English.
XX
XX The present invention relates to novel DNA and protein sequences which
XX are associated with carcinomas. The sequences are useful for: (i) for
XX screening drug candidates; (ii) for screening of bioactive agent capable
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX a bioactive agent capable of modulating the activity of CAP; (iv) for
XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a blichip;
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX carcinoma including lymphoma. The present sequence is one such CA coding
XX sequence. Note: This patent is an equivalent to basic patent
XX US200218286A1, for which no sequence data was published
XX
XX Sequence 38977 BP; 10059 A; 9764 C; 9757 G; 9356 T; 0 U; 41 Other;
XX
XX Query Match 16.5%; Score 30.6; DB 11; Length 38977;
XX Best Local Similarity 26.8%; Pred. No. 24;
XX Matches 40; Conservative 35; Mismatches 74; Indels 0; Gaps 0;
XX
XX QY 6 UAUGGAGUCUUCUUCUUAUAAAUUCUGAUAAGUGCAAGAUUGCUAUAAGU 65
XX Db 30151 TCTTGTGCTGCTTTTAAATCTTAAATTTCTTACAAAAAATTCATTTTGTGATAGT 30210
XX
XX QY 66 AAGGUGCGCUCUUCUUAUAGGCAUACUGGAGUGCGGUGGAGUCCAAAGA 125
XX Db 30211 CTTAATTTGGTTTATTTTAAATGATTTTGTGATTTTCTTAAAGTCAAAAC 30270
XX
XX QY 126 UCCAGGAGCUGAUAUAUUUUUCCUAUAC 154
XX Db 30271 AGCAGGACACTGAACAGCAGATCCCCCCAC 30299
XX
XX RESULT 15
XX ADK53988/c
XX ID ADK53988 standard; DNA; 667 BP.
XX
XX AC ADK53988;
XX
XX 06-MAY-2004 (first entry)
XX
XX Plant DNA sequence which confers altered metabolic characteristic #1371.
XX
XX altered metabolic characteristic; plant; acid metabolism;
XX alcohol metabolism; fatty acid metabolism;
XX branched fatty acid metabolism; alkaloid metabolism;
```

```
KW amino acid metabolism; ester metabolism; glyceride metabolism;
KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;
KW terpene metabolism; isoprenoid metabolism; alkene metabolism;
KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; de.
XX
XX Unidentified.
XX
XX WO2003020936-A1.
XX
XX 13-MAR-2003.
XX
XX 30-AUG-2002; 2002WO-US027884.
XX
XX 31-AUG-2001; 2001US-0316471P.
XX
XX (DOMC ) DOW CHEM CO.
XX
XX (DOMC ) DOW AGROSCIENCES LLC.
XX
XX Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ,
XX Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX WPI; 2003-313091/30.
XX
XX Novel genes that confer altered metabolic characteristics in Nicotiana
XX glauca plants, useful for altering the levels of metabolites e.g.
XX acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX
XX Claim 1; SEQ ID NO 1371; 2576bp; English.
XX
XX The invention comprises DNA sequences which confer an altered metabolic
XX characteristic when they are expressed in a plant. The DNA sequences of
XX the invention are useful for producing plants with an altered metabolic
XX characteristic, such as: altered acid metabolism, alcohol metabolism,
XX fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
XX base metabolism, altered amino acid metabolism, altered ester metabolism,
XX altered glyceride metabolism, altered phenolic metabolism, altered
XX carbohydrate metabolism, altered sterol, oxygenated terpene, or
XX isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
XX metabolism, ketone or quinone metabolism. The DNA sequences of the
XX invention may be used to provide disease resistance in a plant and gene
XX shuffling or sexual PCR procedures. The present nucleic acid represents a
XX DNA sequence of the invention.
XX
XX Sequence 667 BP; 220 A; 141 C; 117 G; 189 T; 0 U; 0 Other;
XX
XX Query Match 16.3%; Score 30.4; DB 10; Length 667;
XX Best Local Similarity 33.1%; Pred. No. 5.6;
XX Matches 45; Conservative 25; Mismatches 66; Indels 0; Gaps 0;
XX
XX QY 10 UGAUCUUCUUCUUAUAAAUUCUGAUAAGUGCAAGAUUGCUAUAAGG 69
XX Db 163 TCTTGTGCTGCTTTTAAATCTTAAATTTCTTACAAAAAATTCATTTTGTGATAGT 104
XX
XX QY 70 UUGCGCUCUUCUUAUAGGCAUACUGGAGUGCGGUGGAGUCCAAAGAUA 129
XX Db 103 TGTGTAAGTCAATTAACCTGTTTACGCTACATGCCCATAGGGTTAAATCAAAAGAT 44
XX
XX QY 130 GGGAGCUGAUAUAUU 145
XX Db 43 GGGATTGTACAAAT 28
XX
XX Search completed: December 22, 2005, 08:40:37
XX Job time : 271.794 secs
```


Qy	60	AUAGUUAAGUGGCGGUCUUAUUAGGCANUACUUCUACAGAUUGCGGCGUCAGUCCA	119
Db	148	TATTTTAAATTAATTAAGGTTAACTATTATTTAGTTTACGTTTACAGATGCTTATGACAGCCCC	207
Qy	120	ACAAGAUCCAGG	131
Db	208	ATAATATCCAGG	219
RESULT 2			
COL53454			
LOCUS	COL53454	486 bp	mRNA
DEFINITION	EN02010.5prime Exelixis FlyTag MN08 Bluescript Drosophila melanogaster cDNA clone EN02010 5, mRNA sequence.		
ACCESSION	COL53454		
VERSION	COL53454.1	GI:48907455	
KEYWORDS	EST		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 486)		
AUTHORS	Nakanishi M., Muzong C., Peterson E., Laufer A., Leung W., Platt D. and Swimmer, C.		
TITLE	Exelixis FlyTag EST Project MN08 Library		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: Stapleton, M. BDGP Lawrence Berkeley National Lab One Cyclotron Rd, Berkeley, CA 94720, USA Fax: 510 486 6798 Email: http://www.fruitfly.org/EST/estefruitfly.berkeley.edu Plate: EN:20 row: A column: 10 High quality sequence stop: 412. Location/Qualifiers		
FEATURES			
SOURCE	1..486 /organism="Drosophila melanogaster" /mol_type="mRNA" /db_xref="taxon:7227" /clone="EN02010" /cell_line="mbn2" /clone_11b="Exelixis FlyTag MN08 Bluescript" /note="Vector: pBluescript; Site 1: NotI; Site 2: XhoI; oligodr primed from LPS induced mbn2 cell line."		
ORIGIN			
Query Match	19.6%;	Score 36.4;	DB 7; Length 486;
Best Local Similarity	38.6%;	Fred. No. 1.5;	
Matches	51; Conservative	28; Mismatches	51; Indels 2; Gaps 1;
Qy	2	UGACUAGUGAUCUGUCUUCUGUAUAAACUGUACAUAAGAAGUCGAAAGUAGUCU--	59
Db	66	TTAAGATGTATCTTGCTCTTATACAAATTTTAGAGGTTTATTAAGAAAGAAATGATGC	125
Qy	60	AUAGUUAAGUGGCGGUCUUAUUAGGCADUACUUCUACAGAUUGCGGCGUUGCAUCCA	119
Db	126	TATTTTAAATTAATTAAGGTTAACTATTATTTAGTTTACGTTTACAGATGCCATTATGGACGCC	185
Qy	120	ACAAGAUCCAGG	131
Db	186	ATAATATCCAGG	197
RESULT 3			
COL53765			
LOCUS	COL53765	490 bp	mRNA
DEFINITION	EN02756.5prime Exelixis FlyTag MN08 Bluescript Drosophila melanogaster cDNA clone EN02756 5, mRNA sequence.		
ACCESSION	COL53765		
VERSION	COL53765.1	GI:48907766	
KEYWORDS	EST		

[illegible]

19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

Location/Qualifiers
source

```

1. 316
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D03005607"
/dev_stage="9 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 9 days embryo"
/notes="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGCGCGCGCACTCGAGTTTCTTTTCTTTT 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATTCTCGAGTTTAAATTAATCCCGCCCCCCC 3']. cDNA was cleaved with BamI and XhoI. Vector: a modified Bluescript KS(+) after bulk excision from Lambda FLC I."
```

ORIGIN

```

Query Match      18.5%; Score 34.4; DB 2; Length 316;
Best Local Similarity 29.1%; Pred. No. 5.8;
Matches 50; Conservative 36; Mismatches 86; Indels 0; Gaps 0;

QY 10 UGACUUGUCUUCGUAUAAUUCUGUACUAAAGUAAUUCGUAAUAGUAAAG 69
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 315 TGCTCTCTTTTAACTTTAAATTTCTTTCACAAAATTTCAATTTTGATGCTCTTA 256
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 70 UUGCGCUGCCUUAUUCGUAUACUUCUGAGAGCGCGUGUGCAACAGAUUCA 129
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 255 TTTCGGTTTCACTTTTAATGCACTTTTGTCAGTTTCTTCATTTTAAATCACAACAGCA 196
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 130 GGAACUGUACAGAAUUCUUAUACCCUGAGUGCGGUGUUGAAUUCUAGUU 181
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 195 GGGACTGAACAGCAGTTCCTCCCACTGCTGCGGGGCTCAGGGCCCGAGTT 144
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

RESULT 7

BP040755

LOCUS

BP040755 616 bp mRNA linear EST 19-AUG-2004

DEFINITION

corniculatus var. japonicus cDNA clone MFB093a04_F 3', mRNA sequence.

ACCESSION

BP040755

VERSION

BP040755.1

KEYWORDS

EST.

SOURCE

ORGANISM

Lotus corniculatus var. japonicus (Lotus japonicus)

Lotus corniculatus var. japonicus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Lotaeae;

Lotus.

1 (bases 1 to 616)

Asamizu, B., Nakamura, Y., Sato, S. and Tabata, S.

TITLE

JOURNAL

PUBMED

COMMENT

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

Location/Qualifiers

FEATURES

FEATURES

source

```

1. 616
/organism="Lotus corniculatus var. japonicus"
/mol_type="mRNA"
/isolate="Miyakojima MG-20"
/db_xref="taxon:34305"
/clone="MFB093a04_F"
/tissue_type="flower bud"
/clone_lib="Lotus corniculatus var. japonicus flower bud"
```

ORIGIN

```

Query Match      18.5%; Score 34.4; DB 3; Length 616;
Best Local Similarity 36.2%; Pred. No. 6.8;
Matches 42; Conservative 23; Mismatches 51; Indels 0; Gaps 0;

QY 56 UGUUAUUGUAGUGUGCGUUCGUUAUUAAGCAUACUUCUGAGUGCGGUGUAG 115
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 456 TGCAATGAGAAAGCTTTGATGTTTGTATTTATTAACGACTTTGTAAGAAACAC 515
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 116 UCCAAACAAGUCCAGGAGACUGUACAGAAUUCUUAUACCCUGAGUGGUGGA 171
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 516 TCCAACAAAATCCGTGACTGTTAGTGTTCCTCGACTGCAGTTCAGCTATGAA 571
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

RESULT 8

CE302369/c

LOCUS

CE302369 651 bp DNA linear GSS 26-SEP-2003

DEFINITION

tigr-gss-dog-17000360210705 Dog Library Canis familiaris genomic,

ACCESSION

CE302369

VERSION

CE302369.1

KEYWORDS

GSS.

SOURCE

ORGANISM

Canis familiaris (dog)

Canis familiaris

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;

Canis.

1 (bases 1 to 651)

Kirkness, E.F., Batina, V., Halpern, A.L., Levy, S., Remington, K.,

Rusch, D.B., DeCher, A.L., Pop, M., Wang, W., Fraser, C.M. and

Venter, D.C.

TITLE

JOURNAL

PUBMED

COMMENT

The dog genome: survey sequencing and comparative analysis

Science 301 (5641), 1898-1903 (2003)

14512627

Contact: Kirkness EF

The Institute for Genomic Research

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,

Rockville, MD 20850, USA

Tel: 301-838-0200

Fax: 301-838-0208

Email: ekirknes@tigr.org

Class: shotgun.

Location/Qualifiers

source

1. 651

/organism="Canis familiaris"

/mol_type="genomic DNA"

/strain="Standard Poodle"

/db_xref="taxon:9615"

/clone_lib="Dog Library"

/note="Site 1: BstXI; Libraries were prepared from

peripheral blood"

ORIGIN

```

Query Match      18.4%; Score 34.2; DB 10; Length 651;
Best Local Similarity 34.7%; Pred. No. 8;
Matches 33; Conservative 24; Mismatches 38; Indels 0; Gaps 0;

QY 2 UGACUUAUGUAGUUCGUUCGUUAUUAUUCUGUACAUAAAGUAAAGUUAUUGU 61
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 342 TGACTGTGAGGCGCTTGTTGTAAGATTCTCTGAGCCCACTTTGTGATGTGCTAT 283
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 62 AGUUAAGUUGCGCUGUUCUUAUUAUAGCAUUAUUC 96
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 282 AGGTAGCATTTGCTTGAAGCTAGCAACACATATCTTC 248
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```


REFERENCE

AUTHORS Sciencemag; Muroidae; Muridae; Murinae; Rattus.

TITLE 1 (bases 1 to 421)

JOURNAL Angen EST Program.

COMMENT Unpublished (2003)
Contact: Dan Fitzpatrick
Angen, Inc

FEATURES

SOURCE One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel.: 805 447-4881
Plate: 00032 row: h column: 1.

Location/Qualifiers
1..421
 /organism="Rattus norvegicus"
 /mol_type="RNA"
 /db_xref="taxon:10116"
 /cclone="cr2-00032-h1"
 /tissue_type="intestinal, fetal"
 /cclone_id="Colon Rat 2 (10396)"
 /note=Vector: pMOB; Site_1: null; Site_2: null; Colon Rat
 2 Fetal rat intestinal library"

ORIGIN

Query Match 18.1%; Score 33.6; DB 6; Length 421;
Best Local Similarity 38.9%; Pred. No. 11;
Matches 28; Conservative 20; Mismatches 24; Indels 0; Gaps 0;

Qy	1 UGACUAGUGAUCUCUGCUCGUAAUAUUAAAUCUGACAUAUAAAGUCAAGAUAUUGCUA 60
Dd	414 TTGTATTTCTTAATTTCCTTATCCATAACATTTGGCAATAGAGTTGAATATGATTCCTG 355
Qy	61 UACUAAAGCUG 72
Dd	354 TAGTTAATGTTG 343

RESULT 11

CW305721/c LOCUS CW305721/c

DEFINITION 639 bp DNA linear GSS 31-OCT-2004

104_791..11466459..148..36273..014 Sordium methylation filtered library (libid: 104) Sordium bicolor genomic clone 11466459, genomic survey sequence.

ACCSSION CW305721

VERSION GW305721.1 GI:55021909

KEYWORDS GSS.

SOURCE Sordium bicolor (sorghum)

ORGANISM Sordium bicolor

REFERENCE Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Sorghum.

1 (base 1 to 639)

Bedell,U.A., Budiman,M.A., Numborg,A., Citek,R.W., Robbins,D., Jones,J., Flick,E., Rohlfing,T., Fries,J.J., Bradford,K., McMenamy,J., Smith,M., Holeman,H., Roe,B.A., Wiley,G., Korf,I.F., Robinsowicz,P.D., Lakey,N., McCombie,W.R., Jeddeleh,J.A. and Martensen,R.A.

REFERENCE

AUTHORS Sordium genome sequencing by methylation filtration

PUBMED PROS Biol. 3 (1), e13 (2005)

15660154

TITLE Contact: Bedell JA

JOURNAL Orion Genomics, LLC

COMMENT 4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: 791 row: d column: 03
Seq primer: SWfor Forward
Class: methylation filtered
High quality sequence stop: 639.
Location/Qualifiers
1..639
 /organism="Sorghum bicolor"
 /mol_type="genomic DNA"
 /cultivar="Atx623"

This Page Blank (usp1c)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: December 22, 2005, 07:55:17 ; Search time 78.2256 Seconds
(without alignments)
4226.575 Million cell updates/sec

Title: US-10-088-750C-5

Perfect score: 186

Sequence: 1 ugcgaucugauucugucuu.....uggaucuaagugucuc 186

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/1/COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5/COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A/COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B/COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H/COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PC/COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP/COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE/COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	36.4	19.6	1345 3	US-09-270-767-1072 Sequence 1072, Ap
2	36.4	19.6	1345 3	US-09-270-767-1072 Sequence 16354, A
3	30.6	16.5	874 3	US-09-020-033-2 Sequence 2, Appl1
4	30.4	16.3	88950 3	US-09-949-016-17150 Sequence 17150, A
5	30.2	16.2	2960 3	US-08-913-842-3 Sequence 3, Appl1
6	30.2	16.2	2960 3	US-09-368-572-3 Sequence 3, Appl1
7	29.6	15.9	1558 2	US-08-455-550-7 Sequence 7, Appl1
8	29.6	15.9	46718 3	US-09-815-093-3 Sequence 3, Appl1
9	29	15.6	36755 3	US-09-949-016-16994 Sequence 16994, A
10	29	15.6	38059 3	US-09-328-925-4 Sequence 4, Appl1
11	28.6	15.4	601 3	US-09-949-016-170384 Sequence 170384, A
12	28.6	15.4	94142 3	US-09-949-016-16553 Sequence 16553, A
13	28.2	15.2	102520 3	US-09-949-016-17367 Sequence 17367, A
14	28.2	15.2	102520 3	US-09-949-016-12448 Sequence 12448, A
15	28	15.1	601 3	US-09-949-016-105589 Sequence 105589, A
16	28	15.1	601 3	US-09-949-016-105591 Sequence 105591, A
17	28	15.1	601 3	US-09-949-016-105591 Sequence 105591, A
18	28	15.1	825 3	US-09-270-767-2680 Sequence 2680, Ap
19	28	15.1	825 3	US-09-270-767-17962 Sequence 17962, Ap
20	28	15.1	129899 3	US-09-949-016-116784 Sequence 116784, A
21	28	15.1	148156 3	US-09-949-016-11776 Sequence 11776, A
22	27.8	14.9	159 3	US-09-621-976-10900 Sequence 10900, A
23	27.8	14.9	3182 2	US-08-971-395-1 Sequence 1, Appl1
24	27.8	14.9	3183 2	US-08-413-135-1 Sequence 1, Appl1

25	27.8	14.9	25249 3	US-09-949-016-17444 Sequence 17444, A
26	27.8	14.9	63860 3	US-09-949-016-15825 Sequence 15825, A
27	27.8	14.9	118382 3	US-09-949-016-15996 Sequence 15996, A
28	27.8	14.9	118382 3	US-09-949-016-15997 Sequence 15997, A
29	27.6	14.8	601 3	US-09-949-016-87413 Sequence 87413, A
30	27.6	14.8	2669 3	US-09-976-594-679 Sequence 679, App
31	27.6	14.8	4080 3	US-09-016-434-1342 Sequence 1342, Ap
32	27.6	14.8	18554 3	US-09-811-825A-3 Sequence 3, Appl1
33	27.6	14.8	18554 3	US-10-652-164-3 Sequence 3, Appl1
34	27.6	14.8	41594 3	US-09-949-016-17298 Sequence 17298, A
35	27.6	14.8	41684 3	US-09-536-059-1 Sequence 1, Appl1
36	27.6	14.8	107941 3	US-09-949-016-14206 Sequence 14206, A
37	27.6	14.8	109378 3	US-09-949-016-12391 Sequence 12391, A
38	27.4	14.7	83516 3	US-09-949-016-15378 Sequence 15378, A
39	27.4	14.6	5568 3	US-09-270-767-26859 Sequence 26859, A
40	27.2	14.6	601 3	US-09-949-016-23557 Sequence 23557, A
41	27.2	14.6	601 3	US-09-949-016-23558 Sequence 23558, A
42	27.2	14.6	601 3	US-09-949-016-158965 Sequence 158965, A
43	27.2	14.6	601 3	US-09-949-016-158966 Sequence 158966, A
44	27.2	14.6	1179 2	US-08-336-778-3 Sequence 3, Appl1
45	27.2	14.6	1332 3	US-09-710-279-1199 Sequence 1199, Ap

ALIGNMENTS

```
RESULT 1
US-09-270-767-1072
Sequence 1072, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OR INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1072
LENGTH: 1345
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-1072

Query Match      19.6%; Score 36.4; DB 3; Length 1345;
Best Local Similarity 38.6%; Pred. No. 0.039;
Matches 51; Conservative 28; Mismatches 51; Indels 2; Gaps 1;

Qy      2 UGACUAGUGAUCUGUCUUGUAUAAUUCUAGUACUAAAGUCGAAAGUAVUGCU-- 59
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      518 TTAAGATGTATCTTCTCTTATACATTTTGAGAGCTTAAATAGAGAAATGATGTC 577

Qy      60 AUAUUAAGUGUGCGUCUUCUUAUUAAGCAUACUUCUCAGAUAGCGCGUGCAUCCA 119
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      578 TATTTAATATATGAGTTACTTATTTAGTTTACTGTTCAGAGATGCTATTGGAGCCCC 637

Qy      120 ACAAGUCCAGG 131
      | | | | | | | |
Db      638 ATAAATATCCAGG 649

RESULT 2
US-09-270-767-16354
Sequence 16354, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
```



```

; SEQ ID NO 16354
;
;   LENGTH: 1345
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-16354

```

Query Match	19.6%;	Score 36.4;	DB 3;	Length 1345;
Best Local Similarity	38.6%;	Pred. No. 0.039;		
Matches 51;	Conservative 28;	Mismatches 51;	Indels 2;	Gaps 14

QY 2 UGACUUAUGAUUUCUGUUCUUAUAAAUUCUGAACUAAAGUCCAAAUAUUGCU -- 59

Db 518 TTAAGATGATCTTCTCTTATCAATTTTGAGAGTTAAATAGAGAGAAATAGTC 577

QY 60 AUAUGUAAGUUGCGUGUCUAUUUAGGCACUUCUCAGAGAGCGGUGUGAGUCCA 119

Db 578 TATTTAAATATAGGTTAACTATTTAGTTTACGTTACAGATGCGCATATGGACGCCCC 637

Qy	120	ACAAGAUC	CAGG	131
Db	638	ATAATATC	CAGG	649

```

RESULT 3
US-09-020-033-2
/ Sequence 2, Application US/09020033A
/ Patent NO. 6043410
/ GENERAL INFORMATION:
/ APPLICANT: Wilkinson, Jack Q.
/ TITLE OF INVENTION: Strawberry Fruit Promoters for Gene Expression
/ FILE REFERENCE: Strawberry Promoters for Gene Expression
/ CURRENT APPLICATION NUMBER: US/09/020,033A
/ CURRENT FILING DATE: 1998-02-06
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 874
/ TYPE: DNA
/ ORGANISM: Fragaria vesca
/ FEATURE:
/ NAME/KEY: promoter
/ LOCATION: (1)..(874)
/ OTHER INFORMATION: RJ39C promoter
US-09-020-033-2

```

Query Match	16.5%;	Score 30.6;	DB 3;	Length 874;
Best Local Similarity	28.7%;	Pred. No. 2.8;		
Matches 45;	Conservative 33;	Mismatches 79;	Indels 0;	Gaps 0;

QY 5 CAAUGGAGCUCUGCUCUUAUAAACUGACUAUAAACUGAAAGUUAUGGUAUAG 64
 Db 597 CGATGATCTCTGTGTCGACATTAACATTTAAACAATATCCAAATCTATATATGTT 656
 QY 65 UAAAGUGGCGUCGUCUNUUAUGCAUUCUCAGAUUGGCGGUGGUCAGUCACAAG 124
 Db 657 TATAATTAATATTATTTCTTAATATTAATCTTCATGTTTAAATTTGCCAGGTCATCAT 716
 QY 125 AUCACAGGACUCUACAGAAUUUUUCUUAACUCUGAGU 161
 Db 717 CTAAATTACGGGTAGAGAAATCTCATGAACCTGGAAT 753

RESULT 4
 US-09-949-016-17150
 ; Sequence 17150, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CLO01307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14

```

? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSQ for Windows Version 4.0
? SEQ ID NO 17150
? LENGTH: 88950
? TYPE: DNA
? ORGANISM: Human
US-03-949-016-17150

```

Query Match	16.3%	Score 30.4	DB 3	Length 88950
Best Local Similarity	26.4%	Pred. No. 14		
Matches 38	Conservative 35	Mismatches 71	Indels 0	Gaps 0

[illegible]

RESULT 5 -842-3/c
US-08-913-842-3/c
Sequence 3, Application US/08913842
Patent No. 6028250
GENERAL INFORMATION:
APPLICANT: OHBA, Toshiaru
APPLICANT: TAKAHASHI, Shuichi
APPLICANT: ANMA, Yoshiko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshi
TITLE OF INVENTION: PLANT PROMOTER AND METHOD FOR GROWING PLANTS
TITLE OF INVENTION: EXPRESSION USING SAID PROMOTER
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.,
STREET: 419 7th Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913,842
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 07-073043
FILING DATE: 30-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/00777
FILING DATE: 26-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25, 618
REFERENCE/DOCKET NUMBER: OHBA=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528


```

INFORMATION FOR SEQ ID NO: 3 :
SEQUENCE CHARACTERISTICS:
LENGTH: 2960 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-913-842-3

Query Match      16.2% Score 30.2; DB 3; Length 2960;
Base Local Similarity 29.7%; Pred.No. 5.7;
Matches 27, Conservative 26; Mismatches 38; Indels 0; Gaps 0.

Oy      15 UUGUCUUGCUAAUAUAUUCGUACUUAAAAAGCGAAGAUAUUGCUAUUGCUAAGSUUGC 74
       :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      310 TTGTGTACTAGTATATAACGATAGTTTAAAGTCACAAATTGTTGAAGATGCATGGTA 251
       ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Oy      75 CUGCCUAUUAGGCAUACUUCUCAGAUVG 105
       ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      250 ATTGCCCTTTTACATGTTTTTTATGATAG 220
       ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 6
US-09-368-572-3/c
Sequence 3, Application US/09368572
Patent No. 6924097
GENERAL INFORMATION:
APPLICANT: OHBA, Toshinaru
            TAKAHASHI, Shuichi
            ANMA, Yoshiko
            ASADA, Kiyoko
            KATO, Ikunoshin
TITLE OF INVENTION: PLANT PROMOTER AND METHOD FOR GENE
                    EXPRESSION USING SAID PROMOTER
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEWMARK, P.L.L.C.
STREET: 419 7th Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/368,572
FILING DATE: 05-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/913,842
FILING DATE: 30-Sept-1997
APPLICATION NUMBER: JP 07-073043
FILING DATE: 30-MAR-1995
APPLICATION NUMBER: PCT/JP96/00777
FILING DATE: 26-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: OHA=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2960 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

```

US-09-368-572-3
Query Match          16.2%; Score 30.2; DB 3; Length 2960;
Best Local Similarity 29.7%; Pred. No. 5.7;
Matches 27; Conservative 26; Mismatches 38; Indels 0; Gaps 0;

Dy      15 UUGCUCUUGCAUAUAACUCGUAACAUAAGCGAAGAAGUAGUCUAAUGUUAAGUCUGG 74
       :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      310 TTTTGTACTACGTAATACGATATGTTTAATAGTCACCATTTGTTTAAGATGCATTGTA 251
       ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Dy      75 CUUGCUAUUUAGGCAUACUCUCAGGAUGG 105
       ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      250 ATTGCCCTCTTTTACATGTTTTTATGATAG 220
       ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 7
US-08-455-550-7/c
Sequence 7, Application US/08455550
Patent No. 567038
GENERAL INFORMATION:
APPLICANT: MORAKAMI, KAZUO
APPLICANT: UENO, NAOTO
APPLICANT: KATO, YUKIO
TITLE OF INVENTION: XENOPUS LAEVIS BONE MORPHOGENETIC PROTEINS AND USE THEREOF
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dike, Bronatein, Roberts & Cushman
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,550
FILING DATE: 31-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/056,564
FILING DATE: 30-APR-1993
APPLICATION NUMBER: 07/577,992
FILING DATE: 05-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Eisenstein, Ronald I
REGISTRATION NUMBER: 30628
REFERENCE/DOCKET NUMBER: 40302-FWC-DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX: 200291
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1558 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEetical: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-455-550-7

Query Match          15.9%; Score 29.6; DB 2; Length 1558;
Best Local Similarity 42.6%; Pred. No. 7.3;
Matches 29; Conservative 15; Mismatches 24; Indels 0; Gaps 0;

Dy      4 ACUAUGAGAUUCUGUCGUAAUAUAACUCGUAACAUAAGCGAAGAAGUUAAGUCUUAUG 63
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```


Db 1427 AATATATGATCATTTGTGCAAAAACATTACACATAAAGTCATAAATAGTCAAG 1368

Qy 64 UUAAGGUU 71

Db 1367 TGAATGTT 1360

RESULT 8

US-09-816-093-3

/ Sequence 3, Application US/09816093

/ Patent No. 6518055

/ GENERAL INFORMATION:

/ APPLICANT: GAN, Weiniu et al

/ TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

/ FILE REFERENCE: CL001182

/ CURRENT APPLICATION NUMBER: US/09/816,093

/ NUMBER OF SEQ ID NOS: 4

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 3

/ LENGTH: 46718

/ TYPE: DNA

/ ORGANISM: Human

/ FEATURE:

/ NAME/KEY: misc feature

/ LOCATION: (1)-(46718)

/ OTHER INFORMATION: n = A,T,C or G

US-09-816-093-3

Query Match 15.9%; Score 29.6; DB 3; Length 46718;

Best Local Similarity 32.6%; Pred. No. 22;

Matches 43; Conservative 25; Mismatches 64; Indels 0; Gaps 0;

Qy 30 AAUCGCAUAAGGUCGAAAGUAUGCUAUAAGGUCGCUUAGG 89

Db 10091 AATGAGGCGAGAGAGATGCTATGTGTAACCAAGATACCATTTCTTGCAAGT 10150

Qy 90 AAUCUCUCAGAGCGCGGUCGACCAAGUACAGAGACUGACGAUUUUC 149

Db 10151 GGGCTTTGTTATGCTGCTTAGGGCTTAAAGCTCAGTGGTGAGATTATCA 10210

Qy 150 UUAACUCGAGU 161

Db 10211 TTGATGGAAT 10222

RESULT 9

US-09-949-016-16994/c

/ Sequence 16994, Application US/09949016

/ Patent No. 6812339

/ GENERAL INFORMATION:

/ APPLICANT: VENTER, J. Craig et al.

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

/ FILE REFERENCE: CL001307

/ CURRENT APPLICATION NUMBER: US/09/949,016

/ CURRENT FILING DATE: 2000-04-14

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-03

/ PRIOR FILING DATE: 2000-10-03

/ PRIOR FILING DATE: 2000-09-08

/ NUMBER OF SEQ ID NOS: 207012

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 16994

/ LENGTH: 36755

/ TYPE: DNA

/ ORGANISM: Human

US-09-949-016-16994

Query Match 15.6%; Score 29; DB 3; Length 36755;

Best Local Similarity 34.1%; Pred. No. 32;

Matches 29; Conservative 21; Mismatches 35; Indels 0; Gaps 0;

Qy 11 GAUCUUGCUUUGCAUAUAUUCGUACAUAAAAGUCGAAAGUAUUGCUAUAAGUUAGU 70

Db 10043 GATTACTTTTGAAGAAATATCATGTGATTCAAAGTTCTAAATTCGCGGAATGAGAT 9984

Qy 71 UGCGCUUGCCUAUUAAGCAUACUU 95

Db 9983 TCGAAGACCTATTGAGACTTAATT 9959

RESULT 10

US-09-328-925-4/c

/ Sequence 4, Application US/09328925

/ Patent No. 6610906

/ GENERAL INFORMATION:

/ APPLICANT: Kurauchi, Kotoku

/ APPLICANT: Kurauchi, Sumiko

/ TITLE OF INVENTION: Nucleotide Sequences for Gene Regulation and Methods of

/ FILE REFERENCE: UM-03603

/ CURRENT APPLICATION NUMBER: US/09/328,925

/ CURRENT FILING DATE: 1999-06-09

/ NUMBER OF SEQ ID NOS: 84

/ SOFTWARE: PatentIn Ver. 2.0

/ SEQ ID NO 4

/ LENGTH: 38059

/ TYPE: DNA

/ ORGANISM: Homo sapiens

US-09-328-925-4

Query Match 15.6%; Score 29; DB 3; Length 38059;

Best Local Similarity 34.1%; Pred. No. 32;

Matches 29; Conservative 21; Mismatches 35; Indels 0; Gaps 0;

Qy 11 GAUCUUGCUUUGCAUAUAUUCGUACAUAAAAGUCGAAAGUAUUGCUAUAAGUUAGU 70

Db 11029 GATTACTTTTGAAGAAATATCATGTGATTCAAAGTTCTAAATTCGCGGAATGAGAT 10970

Qy 71 UGCGCUUGCCUAUUAAGCAUACUU 95

Db 10969 TCGAAGACCTATTGAGACTTAATT 10945

RESULT 11

US-09-949-016-170384

/ Sequence 170384, Application US/09949016

/ Patent No. 6812339

/ GENERAL INFORMATION:

/ APPLICANT: VENTER, J. Craig et al.

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

/ FILE REFERENCE: CL001307

/ CURRENT APPLICATION NUMBER: US/09/949,016

/ CURRENT FILING DATE: 2000-04-14

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-03

/ PRIOR FILING DATE: 2000-09-08

/ NUMBER OF SEQ ID NOS: 207012

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 170384

/ LENGTH: 601

/ TYPE: DNA

/ ORGANISM: Human

US-09-949-016-170384

Query Match 15.4%; Score 28.6; DB 3; Length 601;

Best Local Similarity 35.7%; Pred. No. 12;


```

:
: NUMBER OF SEQ ID NOS: 207012
:
: SOFTWARE: FASTSEQ for Windows Version 4.0
:
: SEQ ID NO. 105589
:
: LENGTH: 601
:
: TYPE: DNA
:
: ORGANISM: Human
:
: OS-09-949-016-105589

```

Query Match 15.1%; Score 28; DB 3; Length 601;
Best Local Similarity 41.3%; Pred. No. 18;
Matches 38; Conservative 14; Mismatches 40; Indels 0; Gaps 0;

[illegible]

Search completed: December 22, 2005, 15:30:24
Job time : 80.2256 secs

Db 214 GGGTTGGCGAGT 203

RESULT 6
US-10-363-483A-19643
Sequence, 19643, Application US/10363483A.
Publication No. US20050064401A1
GENERAL INFORMATION:
APPLICANT: Alexander Olek
APPLICANT: Christian Piepenbrock
APPLICANT: Kurt Berlin
TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
TITLE OF INVENTION: illnesses
FILE REFERENCE: 82011
CURRENT APPLICATION NUMBER: US/10/363,483A
CURRENT FILING DATE: 2003-03-03
NUMBER OF SEQ ID NOS: 40712
SEQ ID NO 19643
LENGTH: 854
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-363-483A-19643
US-10-363-483A-19643

Query Match	18.1%	Score 33.6	DB 9	Length 854
Best Local Similarity	36.1%	Pred. No. 2.3		
Matches	26	Conservative	24	Indels 0
		Mismatches	0	Gaps 0

```

0y      13  UCUCGCUUCGUAUAAUUCUGUACUAAAAAGUCGAAAGUAUUGCUAUAGUUAAGCUG  72
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
db      581 TGTTTTTTGGTTAAAGATTGGATATATAAGTCGAGAGTTTTTTATGTGTTGGGGCG  640

```

```

RESULT 7
US-10-363-483A-19644/C
; Sequence 19644, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 19644
; LENGTH: 854
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CPG-island No: 19644
; US-10-363-483A-19644

```

Query Match	19.1%	Score 33.6	DB 9	Length 854
Best Local Similarity	36.1%	Pred. No. 2.3		
Matches 26; Conservative	22	Mismatches 24	Indels 0	Gaps 0

[illegible]

```

RESULT 8
US-10-125-968-540
; Sequence 540, Application US/10125968
; Publication No. US20030215805A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Palermo, Adam
; APPLICANT: Wang, Youzhen
; APPLICANT: Steimmann, Kathleen
; APPLICANT: Elias, Josh
; APPLICANT: Mertens, Maureen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-032
; CURRENT APPLICATION NUMBER: US/10/125,968
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,163
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 1417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 540
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 315..323, 352, 411
; OTHER INFORMATION: n = A,T,C or G
US-10-125-968-540

```

Query Match	17.8%	Score 33.2;	DB 6;	Length 419;
Best Local Similarity	34.7%	Pred. 0.2,3;		
Matches 52; Conservative	25;	Mismatches	73;	Indels 0;
				Gaps 0;

[illegible]

```

RESULT 9
US-10-972-079-54259
Sequence 54259, Application US/10972079
Publication NO. US20050153317A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: ROSENFELD, David
APPLICANT: KERR, Richard
APPLICANT: BATES, Stephen
APPLICANT: HOLM, Tom
TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF
TITLE OF INVENTION: LIVESTOCK
FILE REFERENCE: MM11110-2
CURRENT APPLICATION NUMBER: US/10/972,079
CURRENT FILING DATE: 2004-10-22
PRIOR APPLICATION NUMBER: US 60/514,333
PRIOR FILING DATE: 2003-10-24
NUMBER OF SEQ ID NOS: 96631
SOFTWARE: Parentlin version 3.1
SEQ ID NO. 54259
LENGTH: 600
TYPE: DNA

```



```
/ ORGANISM: Chicken 1986894306554_1
US-10-972-079-54259

Query Match      17.4%; Score 32.4; DB 9; Length 600;
Best Local Similarity 30.8%; Pred. No. 4.9;
Matches 45; Conservative 30; Mismatches 71; Indels 0; Gaps 0;

QY 1 UUGACCAUGGAGUAGUCUUGCAUUAUUAUUGUACUUAUUAUUGCAUUAUUGCUA 60
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 395 TTCTAAGATTATTAATGCTCAGCTATTAATAATTAATTAATGCTGCTCAGTCTG 454
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 61 UAGUUAAGUUGCGCGUUGCUUUAUUGGCAUACUUCUGGCGGCGUUGCAUCCAA 120
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 455 TGTATCATGTTAATTTTTCAGAAATGCGATTTTTAAATCAATGATGATCTCTGCA 514
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 121 CAAGCAUCCAGGAGCUCUUAACAGAUUU 146
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 515 TAAAGCAATGATGCTACTAGACAGATT 540
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 10
US-09-925-065A-706871/c
/ Sequence 706871, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925, 065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243, 096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252, 147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250, 092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261, 766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289, 846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 706871
/ LENGTH: 1281
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-706871

Query Match      17.3%; Score 32.2; DB 4; Length 1281;
Best Local Similarity 34.0%; Pred. No. 7.8;
Matches 48; Conservative 25; Mismatches 68; Indels 0; Gaps 0;

QY 24 UAAUAAAUCUGUACAUAAAAGUAGUUGCUUUAUUAUUGGCGUUGCUUUAU 83
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 695 TTATTATTAACATTATTAATTAAGAGGCAATGCTATTCCTTGAATTTCTTCGCA 636
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 84 UUAAGCAUACUUCUGGAGUUGCGCGUUGCAUCCAAAGAUUCCAGGAGUUAAGAA 143
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 635 TAAATATCATTTTAAAGAGGAGAGAGAGTTGAAGTGAATTTCAAAATTAATAAAA 576
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 144 UUUUCCUUAUCCUGAGUGCG 164
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 575 TCGTAAGTTAAGTGAAGGTGG 555
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 11
US-09-925-065A-706872/c
/ Sequence 706872, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single

/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925, 065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243, 096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252, 147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250, 092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261, 766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289, 846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 706872
/ LENGTH: 1281
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-706872

Query Match      17.3%; Score 32.2; DB 4; Length 1281;
Best Local Similarity 34.0%; Pred. No. 7.8;
Matches 48; Conservative 25; Mismatches 68; Indels 0; Gaps 0;

QY 24 UAAUAAAUCUGUACAUAAAAGUAGUUGCUUUAUUAUUGGCGUUGCUUUAU 83
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 695 TTATTATTAACATTATTAAGAGGCAATGCTATTCCTTGAATTTCTTCGCA 636
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 84 UUAAGCAUACUUCUGGAGUUGCGCGUUGCAUCCAAAGAUUCCAGGAGUUAAGAA 143
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 635 TAAATATCATTTTAAAGAGGAGAGAGAGTTGAAGTGAATTTCAAAATTAATAAAA 576
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 144 UUUUCCUUAUCCUGAGUGCG 164
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 575 TCGTAAGTTAAGTGAAGGTGG 555
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 12
US-10-425-115-9972
/ Sequence 9972, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425, 115
/ PRIOR FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 9972
/ LENGTH: 2908
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_109093C.1
US-10-425-115-9972

Query Match      17.2%; Score 32; DB 8; Length 2908;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 44; Conservative 9; Mismatches 35; Indels 0; Gaps 0;

QY 93 CUUUCAGAGUUGCGCGUUGCAUCCAAAGAUUCCAGGAGUUAAGAUUUUCCAU 152
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 483 CTCCTCAGGAAGCGGTCTCTCTTCTTAAGATCTCAGGCGCCGATCATCTCTAT 542
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 153 ACCUCGAGUGGAGUUGAUUCCUAGGU 180
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 543 CGCGGAGACCGAGCGGCGCCAAAGT 570
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```



```

RESULT 13
US-10-027-632-174581
; Sequence 174581, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174581
; LENGTH: 713059
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(713059)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-174581

Query Match 17.2%; Score 32; DB 5; Length 713059;
Best Local Similarity 37.5%; Pred. No. 1.4e+02;
Matches 36; Conservative 20; Mismatches 40; Indels 0; Gaps 0;

Cy 7 AUGGAGCUCUCUUCGUAUAAUUCUGUACAUAAGUGCAAGUAGUCUAGUA 66
Db 513039 AGGUGGCTCTGTTGCACAAATGTATTAACCTTAAGGCAATGATTTCTTTAA 513098
Cy 67 AGGUGGCGUCUUCGUAUAAUUCUGUACAUAAGUGCAAGUAGUCUAGUA 102
Db 513099 AAGTCGTTCTTCAAAAGTATGTAATTTTACCTGA 513134

RESULT 14
US-10-027-632-174581
; Sequence 174581, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23

```

```

; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174581
; LENGTH: 713059
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(713059)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-174581

Query Match 17.2%; Score 32; DB 6; Length 713059;
Best Local Similarity 37.5%; Pred. No. 1.4e+02;
Matches 36; Conservative 20; Mismatches 40; Indels 0; Gaps 0;

Cy 7 AUGGAGCUCUCUUCGUAUAAUUCUGUACAUAAGUGCAAGUAGUCUAGUA 66
Db 513039 AGGUGGCTCTGTTGCACAAATGTATTAACCTTAAGGCAATGATTTCTTTAA 513098
Cy 67 AGGUGGCGUCUUCGUAUAAUUCUGUACAUAAGUGCAAGUAGUCUAGUA 102
Db 513099 AAGTCGTTCTTCAAAAGTATGTAATTTTACCTGA 513134

RESULT 15
US-10-425-115-105212/c
; Sequence 105212, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425.115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 105212
; LENGTH: 1038
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_27454C.1
US-10-425-115-105212

Query Match 16.8%; Score 31.2; DB 8; Length 1038;
Best Local Similarity 38.1%; Pred. No. 15;
Matches 32; Conservative 19; Mismatches 33; Indels 0; Gaps 0;

Cy 3 GACUAGUGAUCUCUUCGUAUAAUUCUGUACAUAAGUGCAAGUAGUCUAGUA 62
Db 837 GATTCGTGAAATTCGTTCCAAATTAACCTTGGAACAAGTGAAGATGTTATTACA 778
Cy 63 GUUAGUGGCGUCUUCGUAUAAUUCUGUACAUAAGUGCAAGUAGUCUAGUA 86
Db 777 GATATCTGACAGCTTGATATGTA 754

Search completed: December 22, 2005, 16:21:18
Job time : 439.953 secs

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 08:40:58 : Search time 169.584 Seconds
(without alignments)
569.190 Million cell updates/sec

Title: US-10-088-750C-5
Perfect score: 186
Sequence: 1 uuggaucuaagugacuc.....uuggaucuaagugacuc 186

Scoring table: IDENTITY NUC
Gapop 10.0, 'Gapext 1.0

Searched: 4168288 seqs, 259477437 residues

Total number of hits satisfying chosen parameters: 8336576

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:
1: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31.4	16.9	1713	6	US-10-750-185-55479
2	31	16.7	1825	6	US-10-750-185-33993
3	28.8	15.5	1210	6	US-10-750-185-43180
4	28.6	15.4	187780	6	US-10-995-561-13259
5	28.4	15.3	189252	7	US-11-121-086-54
6	28.2	15.2	207835	7	US-11-121-086-39
7	28.2	15.2	207835	7	US-11-121-086-40
8	27.8	14.9	3183	7	US-11-154-865-1
9	27.6	14.8	2613	6	US-10-821-234-22
10	27.2	14.6	1332	6	US-10-793-626-1199
11	27.2	14.6	2104	6	US-10-793-626-3350
12	27.2	14.6	3153	6	US-10-793-626-3441
13	27.2	14.6	4655	6	US-10-623-155-151
14	27.2	14.6	4849	6	US-10-623-155-335
15	27.2	14.6	160170	7	US-11-121-086-32
16	27.2	14.6	165857	7	US-11-121-086-34
17	27	14.5	2610	6	US-10-750-185-45016
18	27	14.5	101786	7	US-11-117-187-199
19	26.8	14.4	129021	7	US-11-117-187-202
20	26.8	14.4	1021	6	US-10-750-185-55738
21	26.8	14.4	1607	6	US-10-750-185-26224
22	26.8	14.4	3824	6	US-10-750-185-47663
23	26.6	14.3	201	6	US-10-995-561-15796

24	26.6	14.3	2010	6	US-10-750-185-61522	Sequence 61522, A
25	26.6	14.3	2513	6	US-10-750-185-64578	Sequence 64578, A
26	26.6	14.3	2734	6	US-10-510-386-5	Sequence 5, Appl 1
27	26.6	14.3	65723	6	US-10-995-561-13200	Sequence 13200, A
28	26.6	14.3	319608	7	US-11-145-703-1	Sequence 1, Appl 1
29	26.4	14.2	984	6	US-10-750-185-29534	Sequence 29534, A
30	26.4	14.2	1286	6	US-10-750-185-59057	Sequence 62661, A
31	26.4	14.2	1685	6	US-10-750-185-59057	Sequence 59057, A
32	26.2	14.1	1362	6	US-10-750-185-61391	Sequence 61391, A
33	26.2	14.1	1931	6	US-10-750-185-64161	Sequence 64161, A
34	26.2	14.1	2546	6	US-10-750-185-51825	Sequence 31825, A
35	26.2	14.1	5873	6	US-10-750-185-45477	Sequence 45477, A
36	26	14.0	1200	6	US-10-750-185-38659	Sequence 38659, A
37	26	14.0	1392	6	US-10-750-185-61396	Sequence 61396, A
38	26	14.0	2076	6	US-10-750-185-53949	Sequence 53949, A
39	26	14.0	2231	6	US-10-750-185-469	Sequence 469, App
40	25.8	13.9	600	6	US-10-508-263-61	Sequence 61, Appl
41	25.8	13.9	1419	6	US-10-508-263-63	Sequence 11, Appl
42	25.8	13.9	1419	6	US-10-508-263-63	Sequence 11, Appl
43	25.8	13.9	1467	6	US-10-508-263-9	Sequence 9, Appl 1
44	25.8	13.9	1473	6	US-10-508-263-9	Sequence 9, Appl 1
45	25.8	13.9	1545	6	US-10-750-185-53618	Sequence 53618, A

ALIGNMENTS

```

RESULT 1
US-10-750-185-55479
; Sequence 55479, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750.185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55479
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Bovine 19866880687880
US-10-750-185-55479
Query Match 16.9%, Score 31.4, DB 6, Length 1713,
Best Local Similarity 32.7%, Pred. No. 0.39,
Matches 37, Conservative 25, Mismatches 51, Indels 0, Gaps 0;
OY 1 UUGACUAGUCAGUCUUCUGAUAUAAUUCUGUACAUAAAGUGCAAGUAGUCUA 60
DB 617 TTAATCCGTGCTTACGTTCCCTTACATGAAATCAATTAAAGGCCACTACTTATGA 676
OY 61 UAGUAGUGGCGGCGUUCUUAUAGGCAUACUUCUAGGAGGCGGCGUC 113
DB 677 TACACAGATTCTTTTGTATTATTAAGCATTGAAATACAGTGGCTTTTTC 729
RESULT 2
US-10-750-185-33993/c
; Sequence 33993, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.

```



```

/ APPLICANT: KERR, Richard
/ APPLICANT: ROSENFELD, David
/ APPLICANT: HOLM, Tom
/ APPLICANT: BATES, Stephen
/ APPLICANT: PANTIN, Denis
/ TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
/ FILE REFERENCE: M911100-2
/ CURRENT APPLICATION NUMBER: US/10/750,165
/ CURRENT FILING DATE: 2003-12-31
/ PRIOR APPLICATION NUMBER: US 60/437,482
/ PRIOR FILING DATE: 2002-12-31
/ NUMBER OF SEQ ID NOS: 64922
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 33993
/ LENGTH: 1825
/ TYPE: DNA
/ ORGANISM: Bovine
/ OS-10-750-185-33993
198668801131067

```

Query Match	16.7%;	Score 31;	DB 6;	Length 1825;
Best Local Similarity	30.5%;	Pred. No. 0.56;		
Matches 29;	Conservative 26;	Mismatches 40;	Indels 0;	Gaps 0

[illegible]

```

US-10-750-185-43180
RESULT 3
US-10-750-185-43180
Sequence 43180, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMT GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 43180
LENGTH: 1210
TYPE: DNA
ORGANISM: Bovine
US-10-750-185-43180
19866881584705

```

Query Match	15.5%	Score 28.8	DB 6	Length 1210
Best Local Similarity	33.3%	Pred. No. 2,7		
Matches	32	Conservative	22	Mismatches 42
			Indels 0	Gaps 0

2	UGACUAGUAGUAGUCUAGUCGUAUAUAUUCUGACAUAAAAGUCGAAAGUAUUCU	61
1044	TEMAATGAGATTAACCTTCCTTGAAATATGGAAATCAATAAATCAACATTCAAGAC	1103
62	AGUUAAGUGGCGUCCUUAUUAGCAUACUCU	97
1104	AGTAAATGTAAAGCATGAGCTTCATCAAAAAGTTCT	1139

RESULT 4
PS-10-995-561-13259

```

Sequence 13259, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: C1001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ. ID NOS.: 85702
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 13259
LENGTH: 387780
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1) - (387780)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-3)
US-10-995-561-13259

```

Query Match	15.4%	Score 28.6;	DB 6;	Length 387780;
Best Local Similarity	31.3%	Pred. No. 42;		
Matches 26;	Conservative 23;	Mismatches 34;	Indels 0;	Gaps 0;

Dy 6 TAAGCAATCTGCGTAATAAAATCTGACAAUAAAAGCGAAGAATGTCUAATGCU 65
::: :::: : :::: : :::: : :::: :
256036 TATATTAGTCTTTCAAATAATAAAAATATAATTGAAGTAGAAGTTTTCACTTCCTT 256095

Qy 66 AAGGUGCGCCUGCCUAUUG 88
 Db 256096 AAGATTAGTTTGTATTGAAG 256118

```

RESULT 5
US-11-121-086-54
; Sequence 54, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 54
; LENGTH: 189252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-54

```

Query Match	15.3%	Score 28.4	DB 7	Length 189252
Best Local Similarity	25.3%	Pred. NO. 35		
Matches	44	Conservative 39	Mismatches 91	Indels 0
			Gaps 0	
QY	13	UCUCGCUUUCGUAUAAAUUCGUACAUAUAAAGUAGUAUUCGUAGUAAAGUGUG	72	
Db	34931	TTTGGCTATTGTTAAACAGTGTCTCTTAATATGCATCTTTTGGGTTAATTTCAGTAA	34930	
QY	73	CGCUGGCCAUUUUAGCAUACUUCUCAGAGUGCGCGGUCAGUCCAAACAGAUCCAGG	132	
Db	34991	TACATGTAGAGATGAATGTTTTCCTATGATGTATTATCTTCACTTTAAATGGA	35050	
QY	133	ACUGAACAGAAUUUUUCUAUACUCGAGUGCGGUGUGAUAUAAAGUGUACUC	186	
Db	35051	ATGGCAAGACTCTTTTCTTTCTTTTCTTTTGTATTTTAAACGAGATTCACCTC	35104	


```

RESULT 6
US-11-121-086-39/c
; Sequence 39, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: NIELSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: US/11/121, 086
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 39
; LENGTH: 207835
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-39

```

Query Match	Similarity	15.2%	Score 28.2	DB 7	Length 207835
Best Local	Similarity	41.2%	Pred. No. 44		
Matches	Conservative	14	Mismatches	43	Gaps 0

Qy	4	ACUAGUUAUUCUUCGCUUAGUUAUAAAAUCUGUACUAAAAAGUCGAAGUAGUUCUAAUG	63
Db	189015	ACATTGTATTCAGGTTATTCACAAAGTTAAGATGAAGAAAGAAATCTTAAAGCTGTGAGAC	188956

Qy	64	UUAAGGUGGCGGUCGCUAUVUUAAGCAUACUUCUCAG	100
Db	188955	AGAAATCGACGTAACCTTAAAGCAAACTATACG	188919

```

RESULT 7
US-11-121-086-40/c
; Sequence 40, Application US/11121086
; Publication No. US20050266459A1
GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT APPLICATION NUMBER: US/11/121, 086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567, 570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 40
; LENGTH: 207835
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-40

```

	Query Match	15.2%	Score 28.2	DB 7	Length 207835
	Beat Local Similarity	41.2%	Pred. No. 44		
	Matches	40	Conservative	14	Mismatches 43
					Indels 0
					Gaps 0
QY	4	ACUNUGUACUCGUCGUGUAAUAAUAAUCUGACUAAAAGUGGAAAGUAVUGCUAUG	63		
Db	189015	ACATTGTTATCAGGTTATCCAAAGTTTAAATGAATGAAGAAAGAAATCTTAAAGCGCTGTGAGAC	188956		
QY	64	UUAAGUUGCGCGUCGCUAUVUUAAGCAUACUUCUCAG	100		
Db	188955	AGAAAGATCAGGTTAACCTTTAAAGGCAACCTATCAG	188919		

RESULT 8
US-11-154-865-1/c
; Sequence 1, Application US/11154865
; Publication No. US20050251884A1
; GENERAL INFORMATION:

```

? APPLICANT: Gallie, Daniel R.
? APPLICANT: Young, Todd B.
? APPLICANT: The Regents of the University of California
? TITLE OF INVENTION: Generation of Multiple Embryo Maize
? FILE REFERENCE: 023070-121500S
? CURRENT APPLICATION NUMBER: US//11/154,865
? CURRENT FILING DATE: 2005-06-15
? PRIOR APPLICATION NUMBER: US/10/072,077B
? PRIOR FILING DATE: 2002-02-07
? NUMBER OF SEQ ID NOS: 2
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 1
? LENGTH: 3183
? TYPE: DNA
? ORGANISM: Arabidopsis thaliana
? FEATURE:
? NAME/KEY: promoter
? LOCATION: (1)..(3183)
? OTHER INFORMATION: senescence-associated gene 12-1 (SAG12-1) promoter
? US-11-154-865-1

```

	Query Match	14.9%; Score 27.8; DB 7;	Length 3183;
	Best Local Similarity	32.9%; Pred. No. 9.2;	
Matches	Conservative	21; Mismatches	32; Indels
		0; Gaps	0.
Qy	25 AAUAAAUCGUAACUUAAAGUCCAAAUGAUGCUAAGUUAAGUCGCCGUCCUAU	84	
Dd	342 AACCAAGTTTCGTATATAAACCTCGAAATATTGTTTTTAATTATTTGAAGTACATTGT	283	
Qy	85 UAAGCAUACUUCUCAGAUAU	103	
Dd	282 TATGGTTCTTCTTAATGTT	264	

```

RESULT 9
US-10-821-234-22
: Sequence 22, Application US/10821234
: Publication No. US20050255114A1
: GENERAL INFORMATION:
: APPLICANT: Labat, Ivan
: APPLICANT: Stache-Crain, Birgit
: APPLICANT: Andarmani, Susan
: APPLICANT: Tang, Y. Tom
: TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
: FILE REFERENCE: 821A
: CURRENT APPLICATION NUMBER: US/10/821,234
: CURRENT FILING DATE: 2004-04-07
: PRIOR APPLICATION NUMBER: US 60/462,047
: PRIOR FILING DATE: 2003-04-07
: NUMBER OF SEQ ID NOS: 1704
: SOFTWARE: pt_seq_genes Version 1.0
: SEQ ID NO 22
: LENGTH: 2613
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-821-234-22

```

[illegible]

QY 168 UGGAUUCUAGGUGACU 185
| : : ||| : || :
Db 687 CGAGATGTGAGGATGATT 704

[illegible]

NAME/KEY: modified_base	LOCATION: (11838)..(11838)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (11841)..(11841)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (11847)..(11847)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (126041)..(126041)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (126855)..(126855)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (134208)..(134208)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (179735)..(179735)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (93823)..(93823)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (102478)..(102478)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (102481)..(102481)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (103460)..(103460)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (103581)..(103581)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (111433)..(111433)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (115727)..(115727)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (115730)..(115730)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (115734)..(115734)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (120785)..(120785)	OTHER INFORMATION: a, c, g, t, unknown or other
NAME/KEY: modified_base	LOCATION: (120786)..(120786)	OTHER INFORMATION: a, c, g, t, unknown or other


```

; LOCATION: (120788)..(120788)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (120790)..(120790)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (121601)..(121601)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (121603)..(121603)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (121605)..(121605)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (152423)..(152423)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (152427)..(152427)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (152431)..(152432)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (152589)..(152589)
; OTHER INFORMATION: a, c, g, t, unknown or other
; US-11-121-086-32

```

```

Query Match 14.6%; Score 27.2; DB 7; Length 160170;
Best Local Similarity 39.2%; Pred.No.86;
Matches 47; Conservative 15; Mismatches 58; Indels 0; Gaps 0;

```

```

Qy 30 AAUCUGUACUAAAGUCGAAAGUAVUGCUAVUAGUAGUUGCGCUUGCCUAVUUGGC 89
Db 18691 AACCTTGACCATGAAGACAGCACTCAAGGGGTACCCCTGGGCATCCATGA 18632
Qy 90 AUAUCUUCAGAGUAGCGCGUUGCAGUCCACAGAUCCAGGACUAGUAGAATUUUCC 149
Db 18631 ATAAGTCTCCGGCTCCCACTGTAGTCTCTCCAGGCGCCCTTGATATAGTCCAGTCTCTCC 18572

```

```

Search completed: December 22, 2005, 16:41:09
Job time : 172.727 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 04:46:32 ; Search time 1498.86 Seconds
(without alignments)
7205.664 Million cell updates/sec

Title: US-10-088-750C-6

Perfect score: 190
Sequence: 1 cccacaagaugaucuucgu.....cuaaacaagaacaauuaccu 190

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapexc 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_to:*
10: gb_str:*
11: gb_gy:*
12: gb_un:*
13: gb_vl:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	190	100.0	190 6 BD173516	BD173516 Novel ter
2	190	100.0	190 6 BD177020	BD177020 Novel tra
3	190	100.0	8550 13 AF183905	AF183905 Black que
4	63.4	33.4	188 6 BD173514	BD173514 Novel ter
5	63.4	33.4	188 6 BD177018	BD177018 Novel tra
6	63.4	33.4	9185 13 AF218039	AF218039 Cricket p
7	47	24.7	1345 6 AR496112	AR496112 Sequence
8	47	24.7	1345 6 AR511394	AR511394 Sequence
9	45.4	23.9	189 6 BD173513	BD173513 Novel ter
10	45.4	23.9	189 6 BD177017	BD177017 Novel tra
11	45.4	23.9	9264 13 AF014388	AF014388 Drosophila
12	38	20.0	182464 5 BX569779	BX569779 Zebrafish
13	37.6	19.8	30205 5 CR318586	CR318586 Zebrafish
14	37.4	19.7	9275 13 AB017037	AB017037 Himecobi
15	37.4	19.7	9275 13 AB183472	AB183472 Himecobi
16	37	19.5	187 6 BD173512	BD173512 Novel ter
17	37	19.5	187 6 BD177016	BD177016 Novel tra
18	36	18.9	192578 5 AL732455	AL732455 Zebrafish

19	35.6	18.7	152395	8	AL161907	AL161907 Human DNA
20	35.6	18.7	152842	14	AC067921	AC067921 Homo sapi
21	35.6	18.7	181361	14	AC015571	AC015571 Homo sapi
22	35.6	18.7	183466	8	AP005241	AP005241 Homo sapi
23	35.2	18.5	180201	5	BX548064	BX548064 Zebrafish
24	35.2	18.5	200773	9	AL772398	AL772398 Mouse DNA
25	35.2	18.5	226957	14	BX572639	BX572639 Danio rer
26	35	18.4	190708	14	AC132039	AC132039 Ratna no
27	34.8	18.3	69514	8	AY129465_4	AY129465_4 of
28	34.8	18.3	94924	8	AL355997	AL355997 Human DNA
29	34.8	18.3	132200	14	AL161619	AL161619 Homo sapi
30	34.8	18.3	141889	14	AC032006	AC032006 Homo sapi
31	34.8	18.3	148705	14	CR847523	CR847523 Danio rer
32	34.8	18.3	165603	5	BX548250	BX548250 Zebrafish
33	34.8	18.3	167125	14	CR855877	CR855877 Danio rer
34	34.8	18.3	169768	14	BX511118	BX511118 Danio rer
35	34.8	18.3	172764	5	BX294434	BX294434 Zebrafish
36	34.8	18.3	178212	14	AC013326	AC013326 Homo sapi
37	34.8	18.3	231876	5	CR853302	CR853302 Zebrafish
38	34.8	18.3	248999	14	AC120992	AC120992 Ratna no
39	34.8	18.3	316005	14	AC111792	AC111792 Ratna no
40	34.6	18.2	802	10	BV027456	BV027456 S212P024
41	34.6	18.2	163700	8	AC074030	AC074030 Homo sapi
42	34.6	18.2	178453	8	AC016144	AC016144 Homo sapi
43	34.2	18.0	166300	8	AC018884	AC018884 Homo sapi
44	34.2	18.0	185773	14	AC027528	AC027528 Homo sapi
45	34.2	18.0	206214	14	AC079859	AC079859 Homo sapi

ALIGNMENTS

RESULT 1	BD173516	190 bp	RNA	linear	PAT 18-FEB-2003
LOCUS	BD173516				
DEFINITION	Novel tertiary structure having ability to accelerate translation activity.				
ACCESSION	BD173516.1	GI:28414847			
VERSION	BD173516.1				
KEYWORDS	WO 02061080-A/6.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 190)				
AUTHORS	Nakashima, N. and Kanamori, Y.				
TITLE	Novel tertiary structure having ability to accelerate translation activity				
JOURNAL	Patent: WO 02061080-A 6 08-AUG-2002;				
COMMENT	JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA, YASUSHI KANAMORI				
	OS Black queen-cell virus				
	PN WO 02061080-A/6				
	PD 08-AUG-2002				
	PR 31-JAN-2001 WO 2001JP000641				
	PI 25-JAN-2001 JP 01P 016746				
	PC NOBUHIKO NAKASHIMA, YASUSHI KANAMORI				
	CC C12N15/11, C12N15/86, C12P21/02				
	Novel tertiary structure having ability to accelerate translation activity				
	Key translation activity				
	FT source				
	FT source				
FEATURES	Location/Qualifiers				
source	1..190				
	/organism="unidentified"				
	/mol_type="genomic RNA"				
	/db_xref="taxon:32644"				
ORIGIN					
Query Match	100.0%;	Score 190;	DB 6;	Length 190;	
Best Local Similarity	70.5%;	Pred. No. 3e-45;			
Matches 134;	Conservative 56;	Mismatches 0;	Indels 0;	Gaps 0;	

QY 1 CCAACAAGUCGAGUCUCUGCGGAGCGAAGCAAAUUTUGCAGAGUAAAUCUGCAGAGUAGU 60
Db 1 CCAACAAATGTGATCTTGGTCGAGAGCGAAAATTGGACAGATATAAATCTGCAAGTAGT 60
QY 61 GCUAUGUGUGAUAUCCGUAACGCUAUVUUGAGUUNUACGCUACCAAGUUGCGUGUAGUAGC 120
Db 61 GCTATTGTGGAAATCAACCGTACTTATTAGGTTACGGCTCCAGAGTCCGATGAGATGACAGC 120
QY 121 CCUAUCAUAUUCGAGAAACGUGCUAUVUUGUUAAGAAGUAGUAGUCUCUAAAACAGAA 180
Db 121 CCTATCAATATCTAGAGAACTGTCTATGTTAGAGTTAGATGATGCTCTAAACAGAA 180
QY 181 CAUUAUACCU 190
Db 181 CAATTACT 190

RESULT 2
BD177020
LOCUS BD177020 190 bp RNA linear PAT 16-APR-2003
DEFINITION Novel translational activity-promoting higher-order structure.
ACCESSION BD177020
VERSION BD177020.1 GI:30014280
KEYWORDS JP 2002306168-A/6.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 190)
AUTHORS Nakashima, N. and Kanamori, Y.
TITLE Novel translational activity-promoting higher-order structure
JOURNAL Patent: JP 2002306168-A 6 22-OCT-2002;
DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND
ENTOMOLOGICAL HIROSHIMA PREFECTURAL INDUSTRIAL TECHNOLOGY PROMOTION
ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND
FISHERIES
COMMENT OS Black queen-cell virus
PN JP 2002306168-A/6
PD 22-OCT-2002
PF 25-JAN-2001 JP 2001016746
PI NOBHITO NAKASHIMA, YASUSHI KANAMORI
PC C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02// PC
(C12N15/09, C12R1/92), C12N15/00, C12N5/00, (C12N15/00, C12R1/92) CC
Novel translational activity-promoting higher-order structure FH
Key Location/Qualifiers
FT source 1..190
FT /organism="Black queen-cell virus".

FEATURES
source location/Qualifiers
1..190
/organism="unidentified"
/mol_type="genomic RNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 190; DB 6; Length 190;
Best Local Similarity 70.5%; Pred. No. 3e-45;
Matches 134; Conservative 56; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACAAGUCGAGUCUCUGCGGAGCGAAGCAAAUUTUGCAGAGUAAAUCUGCAGAGUAGU 60
Db 1 CCAACAAATGTGATCTTGGTCGAGAGCGAAAATTGGACAGATATAAATCTGCAAGTAGT 60
QY 61 GCUAUGUGUGAUAUCCGUAACGCUAUVUUGAGUUNUACGCUACCAAGUUGCGUGUAGUAGC 120
Db 61 GCTATTGTGGAAATCAACCGTACTTATTAGGTTACGGCTCCAGAGTCCGATGAGATGACAGC 120
QY 121 CCUAUCAUAUUCGAGAAACGUGCUAUVUUGUUAAGAAGUAGUAGUCUCUAAAACAGAA 180
Db 121 CCTATCAATATCTAGAGAACTGTCTATGTTAGAGTTAGATGATGCTCTAAACAGAA 180
QY 181 CAUUAUACCU 190
Db 181 CAATTACT 190

RESULT 3
AF183905
LOCUS AF183905 8550 bp ss-RNA linear VRL 16-AUG-2000
DEFINITION Black queen cell virus nonstructural polyprotein (orf1) and
ACCESSION AF183905
VERSION AF183905.1 GI:8100530
KEYWORDS
SOURCE Black queen cell virus
ORGANISM Black queen cell virus
REFERENCE 1 (bases 1 to 8550)
AUTHORS Leat, N., Bail, B., Govan, V., and Davison, S.
TITLE Analysis of the complete genome sequence of black queen-cell virus,
a picorna-like virus of honey bees
JOURNAL J. Gen. Virol. 81 (Pt 8), 2111-2119 (2000)
FUBMED 10900051
REFERENCE 2 (bases 1 to 8550)
AUTHORS Leat, N. and Davison, S.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-1999) Microbiology, University of Western Cape,
Modderdam Rd., Cape Town, Western Cape 7535, South Africa

FEATURES
source location/Qualifiers
1..8550
/organism="Black queen cell virus"
/mol_type="genomic RNA"
/strain="South African"
/specific_host="Apis mellifera"
/db_xref="taxon:92395"
658..5625
/gene="orf1"
658..5625
/gene="orf1"
/codon_start=1
/product="nonstructural polyprotein"
/protein_id="AAFP2337.1"
/db_xref="GI:8100531"
/translation="MNEPIYENTKVTBILYPPRYMWNKIKLDTLTKEGESVNNK
FLRPTKLRKLCIRLOKDTAEFWKNTSRSLDMIDODIDKVAQSGVMSQSLDRVDM
YHELIHTVVAQNFISIPSHNSKDSKAFVLTQVTSDGMYOINSEYNTATKEVET
FLTKMTTPEQCAEBAHCKIKSPCTLFKCSRHNNOEA.FNFTMSMRLEALAYAPFL
ISRAKTRGASRSTSTYLITLITMTYVDKLVKRLVYYSKHMENAKLAQAGEDS
SFLSVIPLPIESVVDTKSNLFKRIWMSKDLVLVKLSYFGDKITNAFNTGLTWL
TTINKTVAVYKVTLLGHDPBEFLENQSDPLVDMEQCKHFVSDTDRIAITYESTLA
DIKRLVNGVSYMRHPLVYKPHORVISELINQILREADIKKKVGTDAVRNPVLYL
YGEITGVKSTLYPLCATLTKITPREGNTWLSELKOHYKEMIYVRAAEQFPMGYT
QQLVTFDDFNQOVDSANPSLELFEIRSSNIPEYPLHMASIEEKANTVROSXYILC
SSNKTPTKTESLNYKALLRPRAKVEVYKRPASGESCTPCYTVEDPDPHCVIK
SMSRNEILDEVAMTFHEGEPVSSVDKTIENEVFAQGGSGSLSEDDDTLEASVGSN
LDEALIRERKEVILECINDQFNENDTITSFQLIRDFHACADSLQAEFLVARYKSF
DWFSPSKASKVVLGILSLVVGVIYVYVGNPKDASGEYEPVYAKAKVAGENG
PEAYSKYRKINVEGNVARDIKRVEYAVKVNTRVAVESSEVLIHNSLKEKDDI
VQSEACSDINASEQLTAVTNNYVMTLISKEISVVRGHCIFLKGKIAVAALPEPL
ORAYELDGAALIQFNHPYKGNPFETLIDQPSLYKTKNICTKNDLSRDKIAVPVSK
SVYHRDISSYRCELRLOSUGSTRQLDWMKVRVGVITPKVSGSGTSCIKVAVSY
MTDTPDPAEIRLRKAMEYSLETTISGDCGAPLVYNSKIGPEKILGHTAGGHRVGV
SCFATPFYLEDVKEILSLRYDKAQCACELVLRDLBFCALPNHLDEBEFLMGVGRV
PVPQSPRSKISPSPLVGHITERSAPTWLYPRELDGQIFDPLKYRTALGDSVIRS
KLINAKKALIDDIYSVYLQKKDLDDGFSPSLFDDQAVLGIPEGDYNSLKRDPSCG
YVFEKGWTRAKIFGNGDEYDMSTGVOMLEKVOECLBAARGKILDLHYFLDITKDE
RKPKHAKSMFMSGPIIDYIVMSMYNPRIYALVSELKVDHISVGSNSITMDVY
ARYLKSKHNVAGDFEFPDASBQSDIILYAAAGVLOELSKITFNSTEDBMLOQRATIH
SLVNSLHINENGIVLQMKSLPSGHYLAIIINSVFNVLWCLVFNEMAKQKSPYASS
FFREGCIYVAGDDHVSVPKRYLTVFNQOITPLVMSKFGMPTTLETQDTEIDPLSR
LEDVSYLRNFVYDYESRQRYTAPLSLDVLELPMPTKSKDQIVTNVFCNLHAKELSL
LHDKELMEKMSPVYHSKSEVILKMYVSLKPDDEVREIALGSGYE"

gene
CDS
5834..8395
/gene="orf2"
5834..8395
/gene="orf2"

/note="capsid polypeptide; translation initiation most likely facilitated by an internal ribosome entry site (IRES); a region similar to plant scall intestine virus IRES was identified immediately upstream of OR2"

/evidence=not experimental
/transl_except=(pos:5834..5836,aa:Met)
/product="structural polypeptide"
/protein_id="AA072338.1"
/db_xref="GI:8100532"
/translation="MAEIQENYENKQQLVEQETITTFENDLIVLEDPQMEBPLPYA
FHGHTDNRQHTVNFLOPQVIFPDSSWMSLPKPMDSIMIPDIISFPMPEKL
KGFSLRATAVITVQFOTOPQAGRWMSLPPLNTPKPAFNHNSRLMLNHO
HSLASNDTVAVMDKDKRTIMQPSHLAVTSMFKWTGSLVTFKVTGSHGK
PHYFPAQDVAVMDKDKRTIMQPSHLAVTSMFKWTGSLVTFKVTGSHGK
EISFHPSPDYTANSYDVTYRIIVDLREKSESVTIPISFPYKRSRPMDPKSK
YASASTGTIVLKLTSLKATNTVNSVEILIEVAGDDENVIAPIENIPPELSFG
RKGVASQNSGTBOGNRGSLLTDPESITSDPYNPIISLISGFPTNRLIKRY
NFRKATTLNGRISDTPDINSLEAPRLDAQVDTSTKEKGFSTFWSAPTLNIV
AEWALYRGGRVAVKVEKVDFAVATVSPOQYGSVAPATHISTELAEOIKGV
AEFOIPYACLSFRANSETFYSGRNMLDIATSPVNRVYVAGAGDDMFSLF
IGRPCHASQTAQFTKIKQKQVYDLRYDQYDPREVODGTAFIMASIEDSL"

ORIGIN

Query Match 100.0%; Score 190; DB 13; Length 8550;
Best Local Similarity 70.5%; Pred. No. 1.9e-45;
Matches 134; Conservative 56; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACAUUGAUCUGUCGCGAGCAAAUUGCAGAUAAAUCUGCAAGUAGU 60
DB 5647 CCAACAAATGATCTTCTGCGAGGCAAAATTGCACTATTAATCTGCAAGTGT 5706
61 GCUAUGUGAUAUUGCAGUACUUAUUGACGUCUACCAUUGCGUGAUGAGC 120
DB 5707 GCATTTGTGGAAATCAGCTTACCTATTAGGTTACGTCACAAATCGGTGATAGC 5766
QY 121 CCUAUUGAUAUUGAAGACUGUCUUAUUGAAGUAGUAGUCUUAACGAA 180
DB 5767 CCTATCAATATCTGAGGAAGCTGTCTATTAGAGATTAGGTACTCTAAACGAA 5826
QY 181 CAUUAUACU 190
DB 5827 CAATTACT 5836

RESULT 4
BD173514 188 bp RNA linear PAT 18-FEB-2003
LOCUS Novel tertiary structure having ability to accelerate translation activity.
ACCESSION BD173514
VERSION BD173514.1 GI:28414845
KEYWORDS WO 02061080-A/4
SOURCE Cricket paralysis virus
ORGANISM *Cricket paralysis virus*
REFERENCE 1 (bases 1 to 188)
Nakashima, N. and Kanamori, Y.
TITLE Novel tertiary structure having ability to accelerate translation activity
JOURNAL Patent: WO 02061080-A 4 08-AUG-2002;
UPON AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF
SERICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE
FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA, YASUSHI KANAMORI
COMMENT OS Cricket paralysis virus
PN WO 02061080-A/4
PD 08-AUG-2002
PF 31-JAN-2001 WO 2001JP000641
PR 25-JAN-2001 JP 01P 016746

PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI
PC C12N15/11, C12N15/86, C12P21/02
CC Novel tertiary structure having ability to accelerate translation activity
FH Key Location/Qualifiers
FT source 1..188
FT location/Qualifiers
location/Qualifiers
1..188
/organism="Cricket paralysis virus"
/mol_type="genomic RNA"
/db_xref="taxon:12136"

ORIGIN

Query Match 33.4%; Score 63.4; DB 6; Length 188;
Best Local Similarity 42.7%; Pred. No. 7.5e-08;
Matches 79; Conservative 43; Mismatches 61; Indels 2; Gaps 2;

QY 2 CCAACAUUGAUCUGUCGCGAGCAAAUUGCAGAUAAAUCUGCAAGUAGU 61
DB 1 CAAATATGATCTTCTGTTAAATAC-AATTTGAGAGGTTAATTAATCAAGTGTG 59
QY 62 CUUUGUGAUAUUGACCGUACUUAUUGUUAUUGACGUCUACGUGAUGAGCACC 121
DB 60 CTATTTTGTATTAGGTTAGCTATTAGCTTACGTTCCAGATGCTAG-TGCGACCC 118
QY 122 CUUACAUUUGAUAUUGAUAUUGUUAUUGAAGUUAUUGUUAUUGUUAUUG 181
DB 119 CCAACAAATCCAGAGAGCCCTCTGCGGTTTTCAGATTAGGTGCAAAACCTAAG 178
QY 182 AUUU 186
DB 179 AATT 183

RESULT 5
BD177018 188 bp RNA linear PAT 16-APR-2003
LOCUS Novel translational activity-promoting higher-order structure.
ACCESSION BD177018
VERSION BD177018.1 GI:30014278
KEYWORDS JP 2002306168-A/4.
SOURCE Cricket paralysis virus
ORGANISM *Cricket paralysis virus*
REFERENCE 1 (bases 1 to 188)
Nakashima, N. and Kanamori, Y.
TITLE Novel translational activity-promoting higher-order structure
JOURNAL Patent: JP 2002306168-A 4 22-OCT-2002;
DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND
ENTOMOLOGICAL HIROSHIMA PREFECTURAL INDUSTRIAL TECHNOLOGY PROMOTION
ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND
FISHERIES
OS Cricket paralysis virus
PN JP 2002306168-A/4
PD 22-OCT-2002
PF 25-JAN-2001 JP 2001016746
PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI
PC C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02// PC
(C12N15/09, C12R1:92), C12N15/00, C12N5/00, (C12N15/00, C12R1:92) CC
Novel translational activity-promoting higher-order structure FH

COMMENT OS Cricket paralysis virus
PN JP 2002306168-A/4
PD 22-OCT-2002
PF 25-JAN-2001 JP 2001016746
PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI
PC C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02// PC
(C12N15/09, C12R1:92), C12N15/00, C12N5/00, (C12N15/00, C12R1:92) CC
Novel translational activity-promoting higher-order structure FH
Key Location/Qualifiers
FT source 1..188
FT location/Qualifiers
location/Qualifiers
1..188
/organism="Cricket paralysis virus"
/mol_type="genomic RNA"
/db_xref="taxon:12136"

FEATURES

source

ORIGIN

Query Match 33.4%; Score 63.4; DB 6; Length 188;

Best Local Similarity 42.7%; Pred. No. 7.5e-08;
Matches 79; Conservative 43; Mismatches 61; Indels 2; Gaps 2;
QY 2 CAACAUGUAGUCUGUCGAGGCAAAAUUUGCAGUAUAAUUGCAGUAUUG 61
DB 1 CAAAATGAGATCTTGTTAAATAC-AATTTAGAGGTAAATTAATTAACAAGTAC 59
QY 62 CUAAUGUGAUAUCCGUAUUAUUGUUAUUGUUAUUGUUAUUGUUAUUG 121
DB 60 CTAATTTGATTTAGGTTAGCTATTAGCTTTAGCTTCCAGGATGCTAG-TGGCAGCC 118
QY 122 CUAAUUAUUAUUGAUAUUGUUAUUGAUAUUGAUAUUGUUAUUGAUAU 181
DB 119 CCACAATATCCAGAGAACCCCTCTCTGCGGTTTTCAGATTAGTAGTGAAGAAACCTAAG 178
QY 182 AAUUD 186
DB 179 AAATT 183

RESULT 6 9185 bp RNA linear VRL 02-JUL-2000
AF218039
LOCUS Cricket paralysis virus nonstructural polyprotein and structural
DEFINITION polyprotein genes, complete cds.
ACCESSION AF218039.1 GI:8895506
VERSION AF218039.1
KEYWORDS Cricket paralysis virus
SOURCE Cricket paralysis virus
ORGANISM Cricket paralysis virus
Virus; ssRNA positive-strand viruses, no DNA stage;
Dicistroviridae; Crispavirus.
1 (bases 1 to 9185)
Wilson, J.E., Powell, M.J., Hoover, S.E. and Sarnow, P.
Naturally occurring dicistronic cricket paralysis virus RNA is
regulated by two internal ribosome entry sites
Mol. Cell. Biol. 20 (14), 4990-4999 (2000)
10866656
2 (bases 1 to 9185)
Wilson, J.E., Powell, M.J., Hoover, S.E. and Sarnow, P.
Direct Submission
Submitted (20-Dec-1999) Microbiology & Immunology, Stanford
University, 299 Campus Drive, Stanford, CA 94305, USA
Location/Qualifiers
1. 9185
/organism="Cricket paralysis virus"
/virus
/mol_type="genomic RNA"
/db_xref="taxon:12136"
709. 6024
/codon_start=1
/product="nonstructural polyprotein"
/protein_id="AA80998.1"
/translation="MSFOQTNNATNNINSLBELAQLIAQEGNLDGFCFTYVQ
SKPOLDLSEECYCMDPDCGCDIRKREBELRLILTSVYGVNFEWKGWLVKQV
NCPBHRVSTFGNGGLIVSRPFMDHLDFQWCVSSNDECRALFKRTOLEMGD
VSNPBGVQSRPYACNDPRAIRLEALORDEKSTLRIKLRGRIKRRITVGGF
DDLKAGAEVQGNLITRICDFLENSLITLOITVLTITTDKTLNKKDLKRAL
LTVLRLIWMWKKYRAALIVITLFWNPFQDFDQIIDLVDLKDKLITQTAQVETL
SEVFNLSVDEKRVNLGKELGTQGLYDELHWAKEIRHYLDIDENKTLTDETAQ
VEDLVKRLGKVESEKIPDRDIARITTMPLPAGSLVQVLSFGKGGKRPITVWL
TGESGIGKTOMIPLICIDILREMGIVKPAQVIAQVETEVMDGNGOKIVYDL
AFQLODKTQNPETFEVIRTCNTFPOHLMALODKNYSQAEVLVLTNNPOVQLE
SITPDAFYVRMKAHYRVOIKQEKSIWNRAGEVNLADTKLMDKRAIDLSYEP
OKRPFDEATKWDIGDEPISYDEPARTIKANKEKKTFFHOLMLEVASTAAG
GSETSEYVDWDTTFYSNLSQFPAKSLIEAEAFASDAETFNAYITKKNIKET
KSKAMTILDEQISALSTKIRLKNAYKVISHPYLTALGFTGWIISFAMYSFER
TLDDTITSEVSGSDNKTQKSKRYVEVGGSDVTPKPAIVAGVSSGSDSKTMK
KITKVEGSGSDKTOKRNTKVEVKELEKETGCGSDPAHAHLVLDLKNYCL
YVRMVGKWKRYLATATFLRGVCMHPHEFTLRYARKVAPSTNIYSONCDVIT
VIVSHFIAPNARVELTACTRIHDKETPRDCVTVNHRMCHPRDLILGFPYKS

CDS

DOGNLNGFQGLTATFHOSANBELCRAYQWLOAIRPLDQETITVHEDTMDPYESYT
ORDCYENAPOTGNCOSIVGLYNKMERLIGMHIIGNVSECHGAPLTOAIDMG
LNRLEKLDPVNNITVOCCEPSPDIDNNGSGEPBEKFCALIGSKNKGVAATYLLK
SCVYGLSKRITPEALHTRRLPNGITVDPMLGKAGCCGATVDAETVESALDYK
QVVLTOYNSMLDYNKTRRTFTTEAATQGTDDPMKGIAGQISPGRTFQMPRLPQK
QDMWSEGEQYDFESQRAQELRDLRVEBLINDCAAGI1KDVVFDVTLDERARP1EKVAG
KTRVFSAGQHFVVARFRKYFLPFAAYLMMNRIDNEIAVNTVSTDMERLAKLKHG
NKVIADPGNFDLSVAQFQSGCGSPFPMFTFNDVTEDEKRLMICIGLMTIIV
HSVSGDNYVMWTHSPGNGPPTVINCYNMIRIWIILLARLAPEMOSKKER
ENVSMTSYGDNCLNTSDRYEWFNFOIT1SEQMK1KHETDGGKGMVAPPSLSFI
HFLKRFVFSHOCQRTVAPLOQVTEYEMNTFRNTIDPNEILMMNTNFAREIVYHGK
SEVQKRSIGIEDLAKMGILPDQPOLTFRAYLMDATLADDEVDF"
<6217. .8904
/codon_start=1
/product="structural polyprotein"
/protein_id="AA80998.1"
/db_xref="GI:8895508"
/translation="ATFQDKQENSHIENEDKRLMSBQKTI VHFVSGITPSTTALPDI
VNLSTVLDMTTRDRHISIKDFLSGPILITATNLMSSDPEVKQLYANPEVPLISNA
MRODKLGFVGLRATLVKVOVNSOPFOQGRMLQYTPAQVMPNRVTLINETLQGRS
GCEPTDLELVGTEVEMRIPYVSPHYVNLITGQGSFSGIYVVYVQLHDQVSGTSI
EYVMAHLEDDVQVPTGANIFGNSPNYLSIAERATGDPFETTERKXIMHXTYLR
PARIYQAARKELQKLETNNSPTALGQISEGLTTSIHSIPLVGNIBSTPMISAKADL
ATLFGSKPTQGKIGECCTRGCGPMANPDGMDSHKALSTNEIETEGELAGTSLD
ENDLSRVLSIPNWDRTFKTSDVNTVLMNVSPFKVPYSATITDRFRCTHMGV
ANAFYWRGSIYTFKFTQYHSGRLISF1PYYNTTISTGTPDVSRTOKIIVDLR
TSTEVSPVYIASRPMVLCIRBESBWSLKNKDGLAMNCGVIGVRLVNLQVLAQ
NVSEIDVIVCEVSGPDLEFAGPTCPSPYVAGDILLADTRKIEARTOBYENEDNR
ITTCGRIYAOVWGEDQOIPRNEAOGHVP1SIDTHRI5NNMSPOAMCIGKIVSIR
LIRFGI1FGDANTLODSSFVAAPPTVLSITPTNLSRNPQOPDYVLYVAPMGSM
RIKMAVEQDGTGCTPRKNTNFWPFRMFLSDSRSLISTSSAVTTTVLVSCTTMM
GSTQVITDTVSGLEVEVPYINSHITPAVITIDGTSEMDYLKSHSPCLLTSPR
DISATNHHITASFWRAPGDDFSFMYLLEVPVAVARA"

ORIGIN

Query Match 33.4%; Score 63.4; DB 13; Length 9185;
Best Local Similarity 42.7%; Pred. No. 4.6e-08;

Matches 79; Conservative 43; Mismatches 61; Indels 2; Gaps 2;

QY 2 CAACAUGUAGUCUGUCGAGGCAAAAUUUGCAGUAUAAUUGCAGUAUG 61
DB 6029 CAAAATGAGATCTTGTTAAATAC-AATTTAGAGGTAAATTAATTAACAAGTAC 6087
QY 62 CUAAUGUGAUAUCCGUAUUAUUGUUAUUGUUAUUGUUAUUGUUAUUG 121
DB 6088 CTAATTTGATTTAGGTTAGCTATTAGCTTCCAGATGCTAG-TGGCAGCC 6146
QY 122 CUAAUUAUUAUUGAUAUUGUUAUUGAUAUUGAUAUUGUUAUUGAUAU 181
DB 6147 CCACAATATCCAGAGAACCCCTCTCTGCGGTTTTCAGATTAGTAGTGAAGAAACCTAAG 6206
QY 182 AAUUD 186
DB 6207 AAATT 6211
RESULT 7 1345 bp DNA linear PAT 22-SEP-2004
AR496112
LOCUS Sequence 1072 from patent US 6703491.
DEFINITION AR496112
ACCESSION AR496112
VERSION AR496112.1 GI:52431587
KEYWORDS
SOURCE
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 1345)
Homburger, S.A., Ebens, A.J., Jr., Erickson, C.S., Francis-Lang, H.L.,
Margolis, J.S., Reddy, B.P., Ruddy, D.A. and Buchanan, A.R.,
TITLE
Journal Patent: US 6703491-A 1072 09-MAR-2004;
Journal Exelixis, Inc.; South San Francisco, CA
Location/Qualifiers

LOCUS	DEFINITION	LOCUS	DEFINITION
BD173513	Novel tertiary structure having ability to accelerate translation	BD173513	Novel tertiary structure having ability to accelerate translation
189 bp	RNA	189 bp	RNA
linear	PAT 18-FEB-2003	linear	PAT 18-FEB-2003
24.7% Score 47; DB 6; Length 1345;		24.7% Score 47; DB 6; Length 1345;	
37.7% Pred. No. 0.0041;		37.7% Pred. No. 0.0041;	
42; Mismatches 70; Indels 2; Gaps 2;		42; Mismatches 70; Indels 2; Gaps 2;	
Query Match		Query Match	
Best Local Similarity 37.7%;		Best Local Similarity 37.7%;	
Matches 69; Conservative 42;		Matches 69; Conservative 42;	
1. .1345		1. .1345	
/organism="unknown"		/organism="unknown"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity 37.7%;		Best Local Similarity 37.7%;	
Matches 69; Conservative 42;		Matches 69; Conservative 42;	
1. .1345		1. .1345	
/organism="unknown"		/organism="unknown"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity 37.7%;		Best Local Similarity 37.7%;	
Matches 69; Conservative 42;		Matches 69; Conservative 42;	
1. .1345		1. .1345	
/organism="unknown"		/organism="unknown"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity 37.7%;		Best Local Similarity 37.7%;	
Matches 69; Conservative 42;		Matches 69; Conservative 42;	
1. .1345		1. .1345	
/organism="unknown"		/organism="unknown"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity 37.7%;		Best Local Similarity 37.7%;	
Matches 69; Conservative 42;		Matches 69; Conservative 42;	
1. .1345		1. .1345	
/organism="unknown"		/organism="unknown"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity 37.7%;		Best Local Similarity 37.7%;	
Matches 69; Conservative 42;		Matches 69; Conservative 42;	
1. .1345		1. .1345	
/organism="unknown"		/organism="unknown"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity 37.7%;		Best Local Similarity 37.7%;	
Matches 69; Conservative 42;		Matches 69; Conservative 42;	
1. .1345		1. .1345	
/organism="unknown"		/organism="unknown"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity 37.7%;		Best Local Similarity 37.7%;	
Matches 69; Conservative 42;		Matches 69; Conservative 42;	
1. .1345		1. .1345	
/organism="unknown"		/organism="unknown"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity 37.7%;		Best Local Similarity 37.7%;	
Matches 69; Conservative 42;		Matches 69; Conservative 42;	
1. .1345		1. .1345	
/organism="unknown"		/organism="unknown"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity 37.7%;		Best Local Similarity 37.7%;	
Matches 69; Conservative 42;		Matches 69; Conservative 42;	
1. .1345		1. .1345	
/organism="unknown"		/organism="unknown"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity 37.7%;		Best Local Similarity 37.7%;	
Matches 69; Conservative 42;		Matches 69; Conservative 42;	
1. .1345		1. .1345	
/organism="unknown"		/organism="unknown"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity 37.7%;		Best Local Similarity 37.7%;	
Matches 69; Conservative 42;		Matches 69; Conservative 42;	
1. .1345		1. .1345	
/organism="unknown"		/organism="unknown"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
ORIGIN		ORIGIN	
Query Match		Query Match	
Best Local Similarity 37.7%;		Best Local Similarity 37.7%;	
Matches 69; Conservative 42;		Matches 69; Conservative 42;	
1. .1345		1. .1345	
/organism="unknown"		/organism="unknown"	
/mol_type="genomic DNA"		/mol_type="genomic DNA"	
ORIGIN		ORIGIN	

[illegible]

ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND FISHERIES
OS Drosophila C virus
PN JP 2002306168-A/3
PD 22-OCT-2002
PF 25-JAN-2001 JP 2001016746
PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI
PC C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02// PC
(C12N15/09, C12R1/92), C12N15/00, C12N5/00, (C12N15/00, C12R1/92) CC
Novel translational activity-Promoting higher-order structure PH
Key Location/Qualifiers
FT source 1. .189
Location/Qualifiers
1. .189
/organism="Drosophila C virus"
/mol_type="genomic RNA"
/db_xref="taxon:64279"

ORIGIN
Query Match 23.9%; Score 45.4; DB 6; Length 189;
Best Local Similarity 37.7%; Pred. No. 0.016;
Matches 69; Conservative 41; Mismatches 71; Indels 2; Gaps 2;

Qy 7 AUGAUCUCUGUCGAGGCAAAUUTUGCAGAGUAAAUCUGCAAGUAGUCUAAU 66
Db 7 ATGTGATCTTGCTCTCTTATAC-AATTTTGAGAGTTAATAAGAAAGTAGTCTATC 65
Qy 67 GUGGAAUCCAGCCUACUUAUUGGCUUACGCCUCCAGAUCCGUGGAGUAGCCUUAUC 126
Db 66 TTAATTAATTAAGTTAACTTAATTAGTTTACGTTCCAGATGCCCTAT-TGGCAGCCCATTA 124
Qy 127 AAUUAUCUGAGAAACUGUCUUAUUGUUGAAGUAGUAGUCUUAAGCAACAUAUU 186
Db 125 ATATCCAGAGAACCTCTCTCTCTTATATGATTAAGTTAGTTGCAATTGAATTAAGAAAT 184
Qy 187 ACC 189
Db 185 AAC 187

RESULT 11
LOCUS AF014388 9264 bp RNA linear VRL 03-FEB-1998
DEFINITION Drosophila C virus strain EB, complete genome.
ACCESSION AF014388
VERSION AF014388.1 GI:2388672
KEYWORDS
SOURCE
ORGANISM
Drosophila C virus
Drosophila C virus
Viruses; ssRNA positive-strand viruses, no DNA stage;
Dicistroviridae; Crispavirus.
1 (bases 1 to 9264)
REFERENCE
AUTHORS Johnson, K.N. and Christian, P. D.
TITLE The novel genome organization of the insect picorna-like virus Drosophila C virus suggests this virus belongs to a previously undescribed virus family
J. Gen. Virol. 79 (Pt 1), 191-203 (1998)
9460942
2 (bases 1 to 9264)
REFERENCE
AUTHORS Johnson, K.N. and Christian, P. D.
TITLE Direct Submision
J. Gen. Virol. 79 (Pt 1), 191-203 (1998)
9460942
FEATURES
source
1. .9264
/organism="Drosophila C virus"
/mol_type="genomic RNA"
/strain="EB"
/db_xref="taxon:64279"
799. .6078
/note="includes putative protease, helicase and RdRp domains; putative; ORF-1"

/codon_start=1
/product="replicase polyprotein"
/protein_id="AAC58807.1"
/db_xref="GI:2388672"
/translation="MESDKSMAQLNRLINAKMFVEDKSTLTKMADVQKEVYKDPD
AVSPRPAPVFCCTCPFLGTYLMTGIGIKGNKQFEARQMLLLSGDVENPBPVGR
PVYRPNDDPVTYRLERAKIERRDMDIKTLIKELRQIKNRKIYSQGMFDTKQISDGI
KDDVSGEOMNGNLTICDPLENTLPGLQANI QATVTDITDKYSLKEDIMKTLIVLL
VRLMMWKKRRLCYLILFIFKPYGPDQLIDILMDLNKIKISQGLAEITVEVYVH
PMHTCGKILFAMAFLLTKIKPGKDMSPYITRLRIKSIIEGAKITIDYCEVNI
ANQIIMMVLGKTEKTELORANGCYGRIQMAAGROVYIELDQNKDIDLTETANRQI
LWTKLKPSEPLISKEMSLVHTTLLPAKQLETYSCSVKGGFRMPICLMLVGE
SGVKEBMYPLCLIDVLEKMGKIKKODPHQYGVREVEFMGYGQKIVYDDAFQ
KQDKTKANPELFEVIRSCNTEPQHLNMAALHDKNTEKLSKIDKTKAIDLSYEFOK
FDPAPFNRMGMAVYVSPKKEVGIETEKSGSGTKYLKDSKIDKTKAIDLSYEFOK
IVDERKSDGMDIDSGSPDYEDPAKLVCSKMEKAKOSSNNKLFLEVAIRQVSEB
NSEYGCIDPEVDIAKRLQKGETLEIEEDYASDEMPLOYHFKSTIIPASRQCYK
DMDICLSPCKTYLAKYBEIKKILABHPIILITLGMIGALASLAMYTFSSCDEVE
AEVAPSGDAKTYALPRKLVIEIGASGVKQKIVKPEVETEMHNNNGEIEISCEGAM
HRMSAPNNMTDEEFDCTYEDLNKQKRELAQMTSDSWLGEFSLRDRKNKGVIAE
VQSGGVKTNKQIKRVEAGAEELVVALTOGCSDDPAAMIDVFOKTKYMSYPRG
DKAYQLGCTFVAGMSFIMPHYVOAVFARLPNPTIISLOOMSEDLMQIPLSHFS
AGVDNFXLTDNCVRLPFGKNGDFRDCVMVLSMCTPHRDLVRAHPTLSDOCKLFS
SGMAFEHNNMGLYVVMVLAIVRCDKKIKIIPHEDEFEVPEESYIARDCTERYAP
TRDGDGSIIGLKNKILBERKITGMHAGDAEHRGACPLTGCELTARSAVLNKKK
NISOFTYEIPNNVDPDGSVPEGEFVALDGSIRVQAVNSIIPSRILYKLSVPT
MKPALKPTILLNNKHNPLLSGLKCGQVTVAVLSDEYASADVCVWALNQNRLN
KTKYORLILYEEAIRGTODEFMCANINRTSGFPYAQMKRNAFGQMGMGGEEDF
TSNYALALRKDEQLEDCASGKISNVI FVDTLKBERDIAKYNVGTKTFVAGPQHF
VVAFRQYLPFPAMLMENRISNEVAGTVNSDMEIRLAKRLKTGSHVITAGDPCNF
GSLVQIILMAITWEIIFVWLKOFIDISEGRRICICIGMSLHVSHHIEDNYVM
MTHSOPSGNPPTVIINCLYNSITMRLSITRVNEKQROPKSKMFEVEVALITYGDN
VANTDKVEMFNQINISVMTKEMREHYDEKQDITVSRKLEDIPLKRPFRSPE
LQHVAPLKEIVIEYEMLNMSRIDEPDELMSINIEFAFEVYVHGKEBYDKLSAVLA
LQHVQELPENPOLITVNOYILHDIEXVALPLDYF"
<6267. .8972
/note="ORF-2; putative; translation start site unknown"
/codon_start=1
/evidence="experimental"
/product="capsid polyprotein"
/protein_id="AAC58808.1"
/db_xref="GI:2388674"
/translation="ANFOTNNNNINENEDRKITSBOKETVHFSEGGVETSTAVPDIVS
ISTDYLSMTTREDRIHTIKDPLSRPIIOTGMSATTAATFOLYANPREVITSNMY
DEKRLGFGVGLATVIVKIVQNSOPROQRIMQVYPAQYNNRNSVUNSTIQSGSC
PRTDIDLSYGTVEVMKIRIVYSPHYTNALITQGSFGALIVLVYSGLRDVGTGSVEY
TVAHLDEDVDVQPIGANI FTGSSSPNFSGLQMSDGRKTEKDLIDWTSKAVNQP
KIFAVASBEITQLKESGTSISGIGQVSBGLSTWSKIPILGNMFTPAWISAQVSIFFK
MLGFSKPTVOGLPCESKLKGQVRANFGADSHLALSAQMEITEKSGLSGSDSM
DLSHVLSIDNFWDRFTNNTDTATSSILMDNYVTPMKIKPYSSTILDRCTHMGVAN
THGVMGSLVTVFEKVTQFPHSGRLSIFIPYVNTTISAGVPDVSRTQKIVYDLRTS
TEVSFTIYVSSRPMWVCIRPASPWICGDMNLMYNAVIGYIRAVETLQVLAANYFQS
IDTIVEVGGPDLTFAAMPBSYIPYSGAFLTLADDAKAKQKEEEDYDNITPEKISNRG
KREVEDANIVAQWGEDIALQKNDQKHQVHPPTIDTHKIDSWSPEAKICPMTSIR
QIKRFGNALNSLINDAPNTLIAPFVQHPPTVVAAPAEPMSLFEYVYFTGALPFGG
KAFKQAVRTNSAETSVKTDITVWNLNVSQDSFNSLINVSTEDYDIKSTGALPAG
TSGFNSMTYIDPEVEGMEFEIPIYNISSHISPATYVYRGESPTTINSVLRGLHPQ
IVAAVPGTIAATTVDVNAQFAPABSDSPMYIVGVPLTVNARE"

ORIGIN
Query Match 23.9%; Score 45.4; DB 13; Length 9264;
Best Local Similarity 37.7%; Pred. No. 0.0096;
Matches 69; Conservative 41; Mismatches 71; Indels 2; Gaps 2;

Qy 7 AUGAUCUCUGUCGAGGCAAAUUTUGCAGAGUAAAUCUGCAAGUAGUCUAAU 66
Db 6084 ATGTGATCTTGCTCTCTTATAC-AATTTTGAGAGTTAATAAGAAAGTAGTCTATC 6142
Qy 67 GUGGAAUCCAGCCUACUUAUUGGCUUACGCCUCCAGAUCCGUGGAGUAGCCUUAUC 126
Db 6143 TTAATTAATTAAGTTAACTTAATTAGTTTACGTTCCAGATGCCCTAT-TGGCAGCCCATTA 6201
Qy 127 AAUUAUCUGAGAAACUGUCUUAUUGUUGAAGUAGUAGUCUUAAGCAACAUAUU 186


```

/db xref="taxon:7955"
/clone="CH211-239022"
/clone_1fb="CHORI-211"

ORIGIN
Query Match          19.8%; Score 37.6; DB 5; Length 30205;
Best Local Similarity 46.1%; Pred. No. 1.7;
Matches 35; Conservative 17; Mismatches 24; Indels 0; Gaps 0;

Cy      114  UAGAGGCCCUUAUCAAUUCUGAGGAACUGGCUAUCUUAUGAAGAUUGAGUGUCUUA 173
      |||||
Db      10942 TAGAGAGTCATCAATCTTTTAAAGAGATCTGATTATTTTATTAGATCAATTATCAATC 10883

Cy      174  AACAGAACAAUUVUACC 189
      |||||
Db      10882 AATCAACAAATCAACCC 10867

RESULT 14
AB017037          9275 bp      RNA      linear      VRL 22-JAN-2000
LOCUS             Himecobi P virus genomic RNA, complete sequence.
DEFINITION        AB017037
ACCESSION         AB017037
VERSION           AB017037.1 GI:3493357
KEYWORDS          nonstructural protein precursor; capsid protein precursor.
SOURCE            Himecobi P virus
ORGANISM          Himecobi P virus
Virus(es); ssRNA positive-strand viruses, no DNA stage;
                  Distoviridae; Crispavirus.
1 (sites)
Nakashima,N., Sasaki,J. and Toriyama,S.
Determining the nucleotide sequence and capsid-coding region of
himecobi P virus: a member of a novel group of RNA viruses that
infect insects
Arch. Virol. 144 (10), 2051-2058 (1999)
10550677
2 (bases 1 to 9275)
Nakashima,N. and Sasaki,J.
Direct Submission
Submitted (20-AUG-1998) Nobuhiko Nakashima, National Institute of
Sericultural and Entomological Science, Department of Insect
Physiology and Behavior; 1-2 Owasht, Tsukuba, Ibaraki 305-8634,
Japan (E-mail:nakaji@nisee.affrc.go.jp, Tel:81-298-38-6109,
Fax:81-298-38-6028)

FEATURES
         source          location/Qualifiers
         CDS
         1..9275
         /organism="Himecobi P virus"
         /mol_type="genomic RNA"
         /specific_host="Laelophax striatellus"
         /db_xref="taxon:81583"
         964..6297
         /codon_start=1
         /product="nonstructural protein precursor"
         /protein_id="BAA32553.1"
         /db_xref="GI:3493358"
         /translation="MSQLQITKSSKYVNVNOKIKRGSHNHPILRATPNFVEDVESD
LKNLNAIRTTDFKNYVVEITLGHODLRPGFKYNDLELVYASMTSMTETHAPD
LVNULRYVNAHPKDDLDLGGENDSNQYTLSQLKEMRALIWGLQGVDPDILTVP
DCQIALYMTFVLVLDKAVCSRRYRTAYALCSILNHVPHLLPYINKULINFELCVF
FRKLWLYMLKRCFLPSDPLELRMDQVDERMVEEKIEVQEVFNVAOSRSDIE
TIVLSNNEVCSMSVERMYAVCECAEQGLFNVSGLINDTGGYIDSVVSRIGVCS
LVYDMLGLEVAASDILVNLKTYVGLVAVALGVAATCGVMIKLPNPFSLVFKS
DESMNVAQEOSGDTWNGEPLMLNMFVSAGALKSKRCMMAITCSIMLPRASGID
TITFTYKSTYTWGCVRIETSKYVLGVDPGVGVSDSHRPVASKLELGEPIKTSNGTFSY
DSATFTSHISLPIKGLNLTQRSFSSFRSDQAITTGMCKLITERSKNIEASVRNP
PVATYIHGSGVGKSLTIVNLVLAASILSKQIPMMKKGKQKILYSRASQEPFWGTYG
QLVATVDDPSQRADSGAGNPNVELFDIVRAANYPPVPLMANISDKASNTFTSKIIIGS
SNLQKQTESLNPALYRPFVYCVSVSKNEKXNDVPTTHFVEDFYQOEYDMKED
LGTIDWEGYIDKCEVLKGRSPVSLDKEIOELNOSTQFQFVPLNVAOEVQNDYVG
FNCNDGEMGTMCMTNLHQPKKQWMLIKKHIVGTIPGSGVYAEAFKKRPLISQYVLA
SKRPRRWLTSLIKERPVTYKDLRLIHVAATVWKPMPVFGYKKLPAKKECELVHYS
ESTDGNVETPTRESTETEPNVKPTKTESSESNVSVKTESYESNVKVPVENSFPSS
GRSGSENNMLASIQLDNNTYIKEVKEQGVSDQNNAAICSKLVTKNMKRIYVENDHSIPL

```

Query Match	19.7%	Score 37.4	DB 13	Length 9275
Best Local Similarity	35.4%	Pred. No. 2.2		
Matches 67	Conservative 34	Mismatches 86	Indels 2	Gaps 1
QY	2	CAACAUGUGAUCUGCGUCGCGAGCGCAAAUUVUGCAGUAUAAUUCUGCAAGUGG 61		
DB	6285	CGAAAAGTGTGATCTGATTAGAGTAAGAAATCTCTGATTATTAATTTAAATACG 6344		
QY	62	CUAUGUGGAAUACCGGUAUAUUUUGUUAGUACGUCUCAAUGUGUGAUGACGCC 121		
DB	6345	CTACATTTTAAAGACCTTAGTTAGTTTACCTTTACGCCCCAGAT--GGGATCGACGCTT 6402		
QY	122	CUAUAUAUUCUAGAGAACUGUGCUAUGUUUAGAAUUGAGUAGUGUCUCAAACAGAC 181		
DB	6403	CTGTCAATATTCAGAGGCACTAGGTGACGCTGTGATTAGTTAGTGACCTTAGGCTAAG 6462		
QY	182	AUUUUACCU 190		
DB	6463	AATTCTACT 6471		

LOCUS AB183472

DEFINITION Himecobi P virus genomic RNA, complete sequence, isolate: Izumo.

ACCESSION AB183472

VERSION AB183472.1

GI:50251148

SOURCE Himecobi P virus

ORGANISM Himecobi P virus

REFERENCE 1 Nakashima, N., Kawahara, N., Omura, T. and Noda, H.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 04:43:02 ; Search time 274.574 Seconds
(without alignments)
4611.840 Million cell updates/sec

Title: US-10-088-750C-6

Perfect score: 190
Sequence: 1 ccacacagugaucuucgu.....cuaacagaacauuaccu 190

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapexc 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	190	100.0	190	8	ABZ80712 Black que
2	190	100.0	202	6	AA150549 Black que
3	63.4	33.4	188	8	ABZ80710 Cricket P
4	63.4	33.4	200	6	AA150547 Cricket P
5	45.4	23.9	189	8	ABZ80709 Drosophila
6	45.4	23.9	201	6	AA150546 Drosophila
7	37	19.5	187	8	ABZ80708 Himecobi
8	37	19.5	199	6	AA150545 Himecobi
9	33	17.4	2701	5	AA116119 Mouse coc
10	31.6	16.6	1185	11	AC127326 Rice abio
11	31.2	16.4	2000	6	ABZ16307 Arabidops
12	31	16.3	5834	10	ADCC0699 Human nov
13	31	16.3	110000	14	ABZ61095 Human LOC
14	30.8	16.2	52661	9	ADA02876 Human DPT
15	30.8	16.2	52661	10	ADB72614 Human DPT
16	30.8	16.2	52661	10	ADCC8355 Mouse Dpt
17	30.8	16.2	52661	12	ADM74471 Human car
18	30.8	16.2	53795	14	ADZ12703 Human can
19	30.8	16.2	74822	6	ABT10752 Human bre

20	30.8	16.2	91507	4	ABL12300	AB112300 Drosophila
21	30.4	16.0	640	12	ADK89945	Adk89945 S agalact
22	30.4	16.0	1179	12	ADK89938	Adk89938 S agalact
23	30.4	16.0	1179	12	ADK89941	Adk89941 S agalact
24	30.4	16.0	1179	12	ADK89947	Adk89947 S agalact
25	30.4	16.0	1179	12	ADK89940	Adk89940 S agalact
26	30.4	16.0	1179	12	ADK89942	Adk89942 S agalact
27	30.4	16.0	1180	12	ADK89939	Adk89939 S agalact
28	30.4	16.0	1180	12	ADK89943	Adk89943 S agalact
29	30.4	16.0	1305	6	ABN67764	ABN67764 Streptoco
30	30.4	16.0	1308	13	ADV85024	Adv85024 Streptoco
31	30.4	16.0	95596	13	ADV87741	Adv87741 Streptoco
32	30.4	16.0	95596	13	ADV78994	Adv78994 Streptoco
33	30.4	16.0	110000	6	ABN71527_03	Continuation (4 of
34	30.4	16.0	110000	6	ABA03041_10	Continuation (11 of
35	30.4	16.0	110000	6	ABA03041_11	Continuation (12 of
36	30.4	16.0	110000	13	ADV81204_03	Continuation (4 of
37	30.2	15.9	1190	13	ADK64012	Adk64012 Cotton CD
38	30.2	15.9	2000	8	ADA71938	Ada71938 Rice gene
39	30.2	15.9	38258	2	AAV17876	AAV17876 Cloned to
40	30.2	15.9	39005	13	ADR20357	Adr20357 Recombina
41	30.2	15.9	39005	13	ADR21486	Adr21486 Xenorhabd
42	30.2	15.9	47286	12	ADQ97674	Adq97674 Mouse can
43	29.6	15.6	241	2	AAK86847	Aak86847 P. squamu
44	29.6	15.6	668	10	ADB51547	Adb51547 Primary x
45	29.6	15.6	668	13	ADV40339	Adv40339 Rat cardi

ALIGNMENTS

RESULT 1
ABZ80712
ID ABZ80712 standard; RNA, 190 BP.
XX
AC ABZ80712;
XX
DT 15-OCT-2003 (first entry)
XX
DE Black queen-cell virus derived pseudoknot sequence.
XX
KW Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;
KW albumen; impurity; higher-order structure; intergenic region; IGR-IRBS;
KW internal ribosome entry site; ss.
XX
OS Black queen-cell virus.
XX
PN WO2003033719-A1.
XX
PD 24-APR-2003.
XX
PF 08-OCT-2002; 2002WO-JP010447.
XX
PR 17-OCT-2001; 2001JP-00319923.
XX
PA (NAG-) NAT INST AGROBIOLOGICAL SCI.
XX
RA (WAKE-) WAKENYAKU KK.
XX
PI Nakashima N, Shibuya N, Nishikawa S;
XX
DR WPI; 2003-403230/38.
XX
PT Cell-free protein synthesis means in wheatgerm system to establish
PT overexpression of target gene with base sequence sustaining translation
PT activity and function promotion, for producing useful proteins.
XX
PS Claim 1; Page 34; 39pp; Japanese.
XX
CC The invention relates to a cell-free protein synthesis system derived
CC from wheatgerm where there is substantial exclusion of wheatgerm embryo
CC albumen impurities. The novel system uses a sequence having a higher-
CC order RNA structure that promotes translation activity. The higher-order
CC sequence is preferably a "pseudoknot", especially derived from a range of

CC viruses (ABZ80707-ABZ80713). This sequence represents the "pseudoknot"
 CC higher-order sequence from the Black queen-cell virus. The sequence is
 CC used in a construct which may also include an intergenic region and
 CC internal ribosome entry site (IGR-IRIS). The method is applicable in
 CC producing useful proteins

Sequence 190 BP, 58 A; 36 C; 40 G; 0 T; 56 U; 0 Other;

Query Match 100.0%; Score 190; DB 8; Length 190;
 Best Local Similarity 100.0%; Pred. No. 3.8e-54;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACAGUGAUCUGUCUGGAGGAGAAAUUUGCAGUAUAUUGCAAGU 60
 DB 1 CCAACAGUGAUCUGUCUGGAGGAGAAAUUUGCAGUAUAUUGCAAGU 60
 QY 61 GCUAUGUGGAAUVCACCGUACCUAUAUUGAUGUUAAGCUGCAAGUCGUGAUAAGCAGC 120
 DB 61 GCUAUGUGGAAUVCACCGUACCUAUAUUGAUGUUAAGCUGCAAGUCGUGAUAAGCAGC 120
 QY 121 CCUAUCAAUUCUAGAGAAACUGUCUAUUGUUAAGAAUUGUAGUCUCUAAACAGAA 180
 DB 121 CCUAUCAAUUCUAGAGAAACUGUCUAUUGUUAAGAAUUGUAGUCUCUAAACAGAA 180
 QY 181 CAUUVUACCU 190
 DB 181 CAUUVUACCU 190

RESULT 2

AL50549
 ID AAL50549 standard; RNA; 202 BP.

AC AAL50549;
 XX
 DT 07-AUG-2003 (revised)
 DT 19-DEC-2002 (first entry)

XX Black queen-cell virus RNA sequence.

XX CrPV-like virus; ss; higher-order structure; drug development;
 KM drug production; translational activity-promoting function;
 KM protein synthesis; structural analysis.

XX Black queen cell virus.

XX WO200261080-A1.

XX 08-AUG-2002.

XX 31-JAN-2001; 2001WO-JP000641.

XX 25-JAN-2001; 2001JP-00016746.

XX (NAG-) NAT INST AGRBIOLOGICAL SCI.

XX Nakashima N, Kanamori Y,

XX WPI; 2002-627482/67.

XX Translational activity-promoting higher-order structure of CrPV-like
 PT viruses for protein translation when suitably initiated, useful in
 PT synthesis of proteins and polypeptides of foreign species for application
 PT in drugs.

XX Claim 1; Fig 1-2; 38pp; Japanese.

XX The invention comprises seven RNA sequences (CrPV-like viruses) which
 CC have a higher-order structure that sustains translational activity-
 CC promoting function. The RNA sequences of the invention are useful in the
 CC synthesis of proteins and polypeptides for application in developing and
 CC producing drugs. The RNA sequences of the invention are also useful in
 CC basic research of protein synthesis and structural analysis by the gene

CC recombinant technique. The present nucleotide represents a Black queen-
 CC cell virus RNA sequence of the invention. (Updated on 07-AUG-2003 to
 CC correct OS field.)

Sequence 202 BP, 63 A; 38 C; 42 G; 0 T; 59 U; 0 Other;

Query Match 100.0%; Score 190; DB 6; Length 202;
 Best Local Similarity 100.0%; Pred. No. 3.8e-54;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAACAGUGAUCUGUCUGGAGGAGAAAUUUGCAGUAUAUUGCAAGU 60
 DB 1 CCAACAGUGAUCUGUCUGGAGGAGAAAUUUGCAGUAUAUUGCAAGU 60
 QY 61 GCUAUGUGGAAUVCACCGUACCUAUAUUGAUGUUAAGCUGCAAGUCGUGAUAAGCAGC 120
 DB 61 GCUAUGUGGAAUVCACCGUACCUAUAUUGAUGUUAAGCUGCAAGUCGUGAUAAGCAGC 120
 QY 121 CCUAUCAAUUCUAGAGAAACUGUCUAUUGUUAAGAAUUGUAGUCUCUAAACAGAA 180
 DB 121 CCUAUCAAUUCUAGAGAAACUGUCUAUUGUUAAGAAUUGUAGUCUCUAAACAGAA 180
 QY 181 CAUUVUACCU 190
 DB 181 CAUUVUACCU 190

RESULT 3

ABZ80710
 ID ABZ80710 standard; RNA; 188 BP.

AC ABZ80710;

XX
 DT 15-OCT-2003 (first entry)

XX Cricket paralysis virus derived pseudoknot sequence.

XX Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;
 KM albumen; impurity; higher-order structure; intergenic region; IGR-IRIS;
 KM internal ribosome entry site; ss.

XX Cricket paralysis virus.

XX WO2003033719-A1.

XX 24-APR-2003.

XX 08-OCT-2002; 2002WO-JP010447.

XX 17-OCT-2001; 2001JP-00319923.

XX (NAG-) NAT INST AGRBIOLOGICAL SCI.

XX (WAKE-) WAKENYAKU KK.

XX Nakashima N, Shibuya N, Nishikawa S,

XX WPI; 2003-403230/38.

XX Cell-free protein synthesis means in wheatgerm system to establish
 PT overexpression of target gene with base sequence sustaining translation
 PT activity and function promotion, for producing useful proteins.

XX Claim 1; Page 33; 39pp; Japanese.

XX The invention relates to a cell-free protein synthesis system derived
 CC from wheatgerm where there is substantial exclusion of wheatgerm embryo
 CC albumen impurities. The novel system uses a sequence having a higher-
 CC order RNA structure that promotes translation activity. The higher-order
 CC sequence is preferably a "pseudoknot", especially derived from a range of
 CC viruses (ABZ80707-ABZ80713). This sequence represents the "pseudoknot"
 CC higher-order sequence from the Cricket paralysis virus. The sequence is
 CC used in a construct which may also include an intergenic region and
 CC internal ribosome entry site (IGR-IRIS). The method is applicable in

Sequence 200 bp; 60 A; 36 C; 37 G; 0 T; 67 U; 0 other;

Query Match	23.9%	Score 45.4	DB 8	Length 189
Best Local Similarity	60.1%	Pred. No. 4.4e-05		

Matches	110;	Conservative	0;	Mismatches	71;	Indels	2;	Gaps	2;
QY	7	AUGGUAUCUUGCUCUGCGAGGCAAAUUTUGCA	CAGUAUAAAAUCGACAAAGUGUCU	66					
Db	7	AUGGUAUCUUGCUCUCUUAUAC-AAUUTUGAGAGGUAUUAAGAAAGUAGUCU	65						
QY	67	GUUGAAUACCCGUAUCUAUUAAGUUAUCGUC	CAAGUCGUGAUAAGACCCU	126					
Db	66	UUAUAUAUAAGUUAACUAUUAAGUUAUCGUC	UCAGAGCCUUA-UAGGAGCCCCUA	124					
QY	127	AUAUUCUAGAAACUGUGCUAUGUUUAAGAUUAGUAGUCU	UAACAAACAAUUTU	186					
Db	125	AUAUCAGACACCCUCUCUGUCUUAUAUGAUUAGUUCU	UAUUAAGAAAU	184					
QY	187	ACC	189						
Db	185	AAC	187						

RESULT 6
AAL50546
ID AAL50546 standard; RNA; 201 BP.

AC AAL50546;

DT 19-DEC-2002 (first entry)

DE Drosophila C virus RNA sequence.

KW CPV-like virus; BS; higher-order

KW protein synthesis; structural analysis.

Drosophila C virus.

PN WO200261080-A1.

PD 08-AUG-2002

PF 31-JAN-2001; 2001WO-JP000641.

PR 25-JAN-2001; 2001JP-00016746.
XX

PA (NAAG-) NAT INST AGROBIOLOGICAL SCI.
XX

PI Nakashima N, Kanamori Y;
 YX

DR WPI; 2002-627482/67.
xx

PT Translational activity-promoting higher-order structure of CrPV-like
PT viruses for protein translation when sublethally infected [useful for]

synthesis of proteins and polypeptides of foreign species for application in drugs

claim 1: Fida 1-2: 38mn: Japanese

The invention comprises seven RNA sequences (crpV-like viRNAs) which have a higher-order structure that sustains translational activity-promoting function. The RNA sequences of the invention are useful in the synthesis of proteins and polypeptides for application in developing and producing drugs. The RNA sequences of the invention are also useful in basic research of protein synthesis and structural analysis by the gene recombinant technique. The present nucleotide represents a Drosophila C virus RNA sequence of the invention.

Sequence 201 BP; 61 A; 33 C; 35 G; 0 T; 72 U; 0 Other;

Query match	23.98;	Score 45.4;	DB 6;	Length 201;
-------------	--------	-------------	-------	-------------

Matches 110; Conservative 0; Mismatches 71; Indels 2; Gaps 2;

7 AUGUGAUCUUCGUCUGCGGAGGCAAAUUTGACAGUUAUAUUCUGCAAGUAGUCUAUU 66

Db	7	AUGUGAUCUUGCCUUAUA-C-AUUTUGAGAGUUAUAAGAAAGUAGUGCUAUC	65
Qy	67	GUUGGAUVCACCGUACCUAUVUAGUUAUCGCUCCAAAGUUGAGUAGCCCUAUC	12
Db	66	UUAAUAAUUGAGUUAACUAAUUAUUGUUAUCUGUUCAGAAUGCCUAA-UUGCAGCCCCAUA	12
Qy	127	AUAUACUAGAGAAACUGUGCUAUGUUAAGAUAUGUAGUCUCUAAACAGAACAUUU	18
Db	125	AUAUCAGACAGACCCUUCUGUCUUAUAUGUUAUGUUGUUAUUAAGAAUAAGAAAU	18
Qy	187	ACC 189	
Db	185	AAC 187	

RESULT 7
ABZ80708
ID ABZ80708 standard; RNA; 187 BP.

AC ABZ80708;

DT 15-OCT-2003 (first entry)

DE Himetobi P virus derived pseudoknot sequence.

KW Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;

KW internal ribosome entry site; ss.

OS Himetobi P virus. xy

PN WO2003033719-A1.

PD 24-APR-2003 .
XY

PF 08-OCT-2002; 2002W0-JP010447.
XX

PR 17-OCT-2001; 2001JP-00319923.
XX
XX

PA (WAKE-) WAKENYAKTI KK
PA (WAKE-) NAT INST AGROBIOLOGICAL SCI.

XX Nakashima N. Shibuya
PI

WPI: 2003-403230/38.

Cell-free protein synthesis means

PT activity and function promotion, for producing useful proteins

PS Claim 1: Page 32: 39pp: Japanese.

The invention relates to a cell-free protein synthesis system derived from wheatgerm where there is substantial exclusion of wheatgerm embryo albumen impurities. The novel system uses a sequence having a higher-order RNA structure that promotes translation activity. The higher-order sequence is preferably a "pseudoknot", especially derived from a range of viruses (AB28/0707-AB28/0713). This sequence represents the "pseudoknot" higher-order sequence from the hinmeto1 P virus. The sequence is used in a construct which may also include an intergenic region and internal ribosome entry site (IGR-IRE5). The method is applicable in producing useful proteins

50 Sequence 187 BP; 53 A; 32 C; 39 G; 0 T; 63 U; 0 Other;

Query Match 19.5%; Score 37; DB 8; Length 187;

Matches 100; Conservative 0; Mismatches 85; Indels 2; Gaps 1;

Db

2 AAAAUGGUGAUCUGAUGAAGAAAUUCCAGGUAAUUAUUUUUAUACUGCU 61

DB 02 ACACUUUUHAGALCCUUAGUUAAUUAGCUUUUACCGCCACAGAU--UUUUUUUACAGCCUUUU 11

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818

Qy 63 UAUUGUGGAUACAACCGAACCUAAUUNAGUUNUAGCUCUACAAGUGUGUGAUNAGACGCC 122
 Db 2316 TGAGGCTTGGTCTCATTACACTACATTCATTTTAAACCAATGATGTGGACAGGTGGC 225
 Qy 123 UAUCAUAUUCUAGAAGAACUGUCUAGUUNUUGAAGAUUAGUGUACUUAACAGAAACA 182
 Db 2356 TTGTATGTATCAAGAGAGTGTGTCTTAAGCTTAAAGGTCCCTGAGAGTATCTGAAGAACATA 2197
 Qy 183 AUUUACCU 190
 Db 2196 CTTGGCT 2189

RESULT 10
 ACCL27326/C
 ID ACCL27326 standard; cDNA; 1185 BP.
 XX
 AC ACCL27326;
 XX
 DT 02-JUN-2005 (first entry)
 XX
 DE Rice abiotic stress responsive polynucleotide SEQ ID NO:1282.
 XX
 KW es; abiotic stress tolerance; transgenic plant; plant; cereal;
 KW agriculture.
 XX
 OS Oryza sativa.
 XX
 PN WO2003008540-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 21-JUN-2002; 2002WO-US019666.
 XX
 PR 22-JUN-2001; 2001US-0300112P.
 PR 24-AUG-2001; 2001US-0314662P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 21-NOV-2001; 2001US-0332132P.
 XX
 PA (SYCN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Krieps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,
 PI Moughamer T, Provart N, Rieke D, Zhu T;
 XX
 DR WPI; 2003-248011/24.
 XX
 PT New stress-responsive nucleic acid, useful for altering the
 PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
 PT stress, salt stress or osmotic stress.
 XX
 PS Claim 1; SEQ ID NO 1282; 89pp; English.
 XX
 CC The invention relates to novel abiotic stress responsive polynucleotides
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 CC cells, and plants containing such polynucleotides. Also disclosed are
 CC methods for using the polynucleotides and polypeptides to alter the
 CC responsiveness of a plant to abiotic stress. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulated
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
 CC molecule and the polypeptide encoded by it are useful in altering the
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
 CC stress, osmotic stress or any of their combinations. The present sequence
 CC is used in the exemplification of the invention
 CC
 CQ Sequence 1185 BP; 366 A; 319 G; 197 G; 303 T; 0 U; 0 Other;

QY 2 CAACAAUGGUAUUCUUCUGCGAGGACAAAAUUCGACAGUAUAAAUCUGCAAGUAGUG 61
 Db 673 CAAAGATATTTGGATCTTCGTAGACATGCAAAAGTTTTTGGGGATATAGGCCCTCACATTTTCAT 614
 QY 62 CUAAUUGUGGAUACACCGUACCUAUUUAGUUUACGCUCCAAAGGCGAGUGAUNAGACGC 121
 Db 613 TACCATTATATGTGCAAAAGTAGCTAGGTTGGTTAGATTGGCTAGAGTTCTGGGATCTCTC 554
 QY 122 CUAUCAUAUUCUAGAGAAUCUGGUCUAUGUUUAGAGUAUUGUAG 167
 Db 553 CGGTCATATGTTGTTGCTTAGACTTAATCACTGACGATTAAGTCAG 508

RESULT 11
 ID ABZ16307/C
 ABZ16307 standard; DNA; 2000 BP.
 AC
 XX ABZ16307;
 XX
 DT 21-JAN-2003 (first entry)
 XX
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 4112.
 XX
 KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200216655-A2.
 XX
 XX 28-FEB-2002.
 PF
 XX 24-AUG-2001; 2001WO-US026685.
 PR 24-AUG-2000; 2000US-0227866P.
 PR 26-JUN-2001; 2001US-0266447P.
 PR 22-JUN-2001; 2001US-0300111P.
 XX
 PA (SCRI) SCRIPPS RES INST.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 PI
 XX Harper JF, Krepe J, Wang X, Zhu T;
 DR WPI; 2002-304127/34.
 XX
 PT Identifying a stress condition to which a plant cell has been exposed and
 PT producing plants with increased tolerance to these abiotic stressors.
 XX
 PS Claim 144; SEQ ID NO 4112; 577bp + Sequence listing; English.
 CC
 CC The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office
 XX
 XX Sequence 2000 BP; 613 A; 306 C; 304 G; 777 T; 0 U; 0 Other;

	Query Match	Similarity	16.4%	Score 31.2	DB 6	Length 2000
	Best Local	Similarity	44.0%	Prod. No. 6.6		
	Matches	37	Conservative	14	Mismatches	33
					Indels	0
					Gaps	0
Oy	97	GCUCCAAGAUCCGUGGAGUACGACGCCUACAUAUUCUAGAGAACTUGCGUAUUNUGA	156			
Db	1975	GTTCCTCGAATCAGATGATATACAGCTCTTTTGGTTACCAAAAGAACTGATATCAGAAA	1916			
Oy	157	AGAUUAGUAGUCUCUUAACAGAA	180			

Db 1915 AAATTCAAGAGCTCTTAATTGAA 1892

RESULT 12
ADCC30699
ID ADCC30699 standard; cDNA; 5834 BP.

AC ADCC30699;
DT 18-DEC-2003 (first entry)
XX Human novel cDNA sequence, SEQ ID NO:781.

XX Human, diagnostic; drug screening; forensics; gene mapping;
KM biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KM neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KM ulcers; osteoporosis; autoimmune disease; cancer;
KM molecular weight marker; food supplement; antiparkinsonian; nootropic;
KM neuroprotective; anti-anemic; anticoagulant; thrombolytic; vulnerrary;
KM antituber; osteopathic; immunosuppressive; anti-inflammatory; cytostatic;
KM gene therapy; gene; ss.

XX Homo sapiens.
OS
XX MO2003029271-A2.
PN
PD 10-APR-2003.
XX
XX 24-SEP-2002; 2002WO-US030474.
PF
XX 24-SEP-2001; 2001US-0324631P.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QH, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Aarndt V, Wang Z, Wang G;
PI Haley-Vicente D, Drmanac RT;
XX
XX WPI; 2003-371981/35.
DR P-PSDB; ADCC31670.
XX

PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemia, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.

PS Claim 1; SEQ ID NO 781; 1185bp; English.

XX The invention relates to 971 novel human cDNA sequences (ADCC3919-
CC ADCC30889) and the polypeptides they encode (ADCC30890-ADCC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADCC31861-ADCC32627) and the polypeptides encoded by the contigs (ADCC32628
CC -ADCC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or

CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 5834 BP; 1332 A; 1589 C; 1494 G; 1419 T; 0 U; 0 Other;

SO Query Match 16.3%; Score 31; DB 10; Length 5834;
Best Local Similarity 31.9%; Pred. No. 11;
Matches 43; Conservative 27; Mismatches 65; Indels 0; Gaps 0;

QY 25 AGGCAAAUUGUCACAGUAAUUCUCCAGUAGUGCUUAGUUGAUAUACCGUACU 84
DB 4117 AGACACATTCTGGGACGTAAATATATATGTTGTTATTTGTAACCTTGTA 4176
QY 85 AUUUGAGUUGACGUCUCCAGUAGUGGUGAUGACGCCUUAUUAUAGAGAAUGU 144
DB 4177 ATTATTAATCTGCTCTTGATTTGATGGGCACTTGAAGTACCTGTGGGTAGAGGATATG 4236
QY 145 GCUAUGUUGAAGA 159
DB 4237 CTTATTTCTGAGGA 4251

RESULT 13
AEA61095_0/c
WP Sequence split into 4 fragments LOCUS AEA61095 Accession Aea61095
WP Fragment Name Begin End
WP AEA61095_0 1 110000
WP AEA61095_1 100001 210000
WP AEA61095_2 200001 310000
WP AEA61095_3 300001 380963
ID AEA61095 standard; DNA; 380963 BP.

XX AEA61095;
XX
XX 25-AUG-2005 (first entry)
DT
XX Human LOC339479 gene genomic sequence SEQ ID NO:5.
DE
XX DNA methylation; biomarker; cancer; gene; de; LOC339479.
KM
XX Homo sapiens.
OS
XX US2005130172-A1.
PN
XX 16-JUN-2005.
PD
XX 27-JAN-2004; 2004US-00765790.
PF
XX 16-DEC-2003; 2003US-00737082.
PR
XX (FARB) BAYER CORP.
PA
XX Beard C, Burgess C, Gannon A, Harvey J, Lechner JF, Li Z;
PI WPI; 2005-456991/46.
DR GENBANK; BF589529.
DR

PT Identifying nucleic acid sequences as biomarker for disease, by
PT identifying nucleic acid sequences comprising methylated CpG site and
PT down-regulated in diseased cells and comparing its expression level with
PT demethylated nucleic acid.

XX Claim 11; SEQ ID NO 5; 27pb; English.

XX The invention relates to a method (M1) for identifying one or more
CC nucleic acid sequences useful as a biomarker for a disease to be
CC detected. (M1) involves identifying nucleic acid sequences comprising

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 07:11:33 ; Search time 1988.22 Seconds
(without alignments)
4471.109 Million cell updates/sec

Title: US-10-088-750C-6

Perfect score: 190
Sequence: 1 ccaacagaugacuuacuu.....cuaacagacacuuuacuu 190

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 2393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:
10: gb_est10:
11: gb_est11:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	24.7	433	7	CO153155 EN01401.5
2	47	24.7	486	7	CO153454 EN02010.5
3	47	24.7	490	7	CO153765 EN02756.5
4	47	24.7	539	7	CO337794 EN15317.5
5	36.8	19.4	714	6	CA248993 SCSEFL110
6	35.8	18.8	826	10	CM675417 OP_BA007
7	35.6	18.7	831	5	BQ41285 AGENCOURT
8	35.4	18.6	751	7	CK121556 SB02035A1
9	34.8	18.3	608	10	BX120183 Danilo rer
10	34.8	18.3	1228	2	BG499554 602546703
11	34.4	18.1	601	11	CPA562935
12	33.8	17.8	574	3	BP912007
13	33.8	17.8	700	5	BX615168
14	33.4	17.6	902	8	CL685543
15	33.2	17.5	402	8	R25253
16	33.2	17.5	593	9	AO680797
17	33.2	17.5	893	10	CZ958530
18	33.2	17.5	952	10	CZ958530
19	33.3	17.4	1101	10	CNS01289
20	33.8	17.3	366	10	CM108676
21	33.8	17.3	547	10	CZ042704
22	32.8	17.3	647	10	CM328338

C 23	32.8	17.3	724	10	CL381173	CL381173 RPT44_31
C 24	32.8	17.3	763	2	BBS69614	BBS69614 601328256
C 25	32.8	17.3	767	10	CM359111	CM359111 fbb0001f0
C 26	32.8	17.3	781	10	CM375429	CM375429 fbb0001f0
C 27	32.8	17.3	977	9	CC217149	CC217149 CC261-23E
C 28	32.6	17.2	451	6	CA433304	CA433304 UI-H-CO0-
C 29	32.6	17.2	759	10	BX207527	BX207527 Danilo rer
C 30	32.6	17.2	822	10	BX221224	BX221224 Danilo rer
C 31	32.4	17.1	514	2	BB703319	BB703319 BB703319
C 32	32.4	17.1	579	2	BB080560	BB080560 H3055A10-
C 33	32.4	17.1	723	10	CZ754700	CZ754700 OC_BA010
C 34	32.4	17.1	883	7	CO116428	CO116428 GR_EB018
C 35	32.4	17.1	899	10	DU074448	DU074448 69473 Tom
C 36	32.2	16.9	668	10	CZ057824	CZ057824 OM_BA005
C 37	32.2	16.9	719	9	CE167648	CE167648 tigr-g88-
C 38	32.2	16.9	727	1	AV707635	AV707635 AV707635
C 39	32	16.8	522	5	BY471748	BY471748 BY471748
C 40	32	16.8	712	7	CK308262	CK308262 SB02046A1
C 41	32	16.8	716	5	BQ781365	BQ781365 UI-R-PF0-
C 42	32	16.8	754	8	DN035261	DN035261 JGI_CAA01
C 43	32	16.8	807	7	CK132357	CK132357 SB02011A2
C 44	32	16.8	842	8	DR660385	DR660385 JGI_CABG4
C 45	32	16.8	849	8	DN087727	DN087727 JGI_CABE1

ALIGNMENTS

RESULT 1
LOCUS CO153155 433 bp mRNA linear EST 05-OCT-2004
DEFINITION EN01401.Sprine Exelixis FlyTag MN08 Bluescript Drosophila
melanogaster cDNA clone EN01401 5, mRNA sequence.

ACCESSION CO153155
VERSION CO153155.1 GI:48907156
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 433)
Nakanishi, M., Muzong, C., Peterson, E., Laufer, A., Leung, W., Platt, D.
and Swimmer, C.
Exelixis FlyTag EST Project MN08 Library
Unpublished (2004)
CONTACT: Stapleton, M.
BDGP

TITLE Exelixis FlyTag EST Project MN08 Library
JOURNAL Unpublished (2004)
COMMENT Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
Plate: EN.14 row: A column: 1
High quality sequence stop: 383.
Location/Qualifiers

FEATURES
source 1..433
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="EN01401"
/cell_line="mdn2"
/clone_lib="Exelixis FlyTag MN08 Bluescript"
/note="Vector: pBluescript; Site 1: NotI; Site 2: XhoI;
oligodt primed from LPS induced mdn2 cell line."

ORIGIN
Query Match 24.7%; Score 47; DB 7; Length 433;
Best Local Similarity 37.7%; Pred. No. 0.0008;
Matches 69; Conservative 42; Mismatches 70; Indels 2; Gaps 2;

Qy 7 AUGGAGUCUCUCGCGGAGCAAAUUGCAGAGUAAUUCGCAAGAGUCUCUUV 66
Db 93 ATGTCATCTTCCTTCCTATAC-AATTTTGAGAGGTTAATAAGAAAGAGTGTCTATT 151

QY 67 GUUGAUAUCCGUAUUAUAGGCUCCAAAGUGGUAUAGCAAGCCCUAUC 126
 DB 152 TTAAATTAATTAAGTTAACTTAAGTTAGTTCTGATGAGTGCCTAT-TGGCAGCCCCATA 210
 QY 127 AAUUAUCUAGAGAAACUGUCUAUUGUUUAGAAGUUAUGUAGUCUUAACAGAACAAU 186
 DB 211 ATATCCAGAGACACCCCTCTGCTCTTATATGATTAAGTTGTCATTAGAAATGAAGAAAT 270
 QY 187 ACC 189
 DB 271 AAC 273

RESULT 2
 CO153454 486 bp mRNA linear EST 05-OCT-2004
 LOCUS EN02010.5prime Exelixis FlyTag MN08 Bluescript Drosophila
 DEFINITION melanogaster cDNA clone EN02010 5, mRNA sequence.
 ACCESSION CO153454
 VERSION CO153454.1 GI:48907455
 KEYWORDS EST
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 486)
 Nakanishi, M., Muzong, C., Peterson, E., Laufer, A., Leung, W., Platt, D.
 and Swimmer, C.
 Exelixis FlyTag EST Project MN08 Library
 Unpublished (2004)
 COMMENT Contact: Stapleton, M.
 BDCP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
 Plate: EN.20 row: A column: 10
 High quality sequence stop: 412.
 Location/Qualifiers
 1..486
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="EN02010"
 /cell_line="mbn2"
 /clone_1lb="Exelixis FlyTag MN08 Bluescript"
 /note="Vector: pBluescript; Site_1: NotI; Site_2: XhoI;
 oligodT primed from LPS induced mbn2 cell line."

ORIGIN
 Query Match 24.7%; Score 47; DB 7; Length 486;
 Best Local Similarity 37.7%; Pred. No. 0.00082;
 Matches 69; Conservative 42; Mismatches 70; Indels 2; Gaps 2;

QY 7 AUGUGAUCUUGCUGCGGAGCAAAUUGCAGUAUAAUUCGCAAGUAGUCUAU 66
 DB 71 ATGTGATCTTGCTTCTTATAC-AATTTTGAGAGGTTAATGAAGAGTAGTGCTATT 129
 QY 67 GUUGAUAUCCGUAUUAUAGGCUCCAAAGUGGUAUAGCAAGCCCUAUC 126
 DB 130 TTAAATTAATTAAGTTAACTTAAGTTAGTTCTGATGAGTGCCTAT-TGGCAGCCCCATA 188
 QY 127 AAUUAUCUAGAGAAACUGUCUAUUGUUUAGAAGUUAUGUAGUCUUAACAGAACAAU 186
 DB 189 ATATCCAGAGACACCCCTCTGCTCTTATATGATTAAGTTGTCATTAGAAATGAAGAAAT 248
 QY 187 ACC 189
 DB 249 AAC 251

RESULT 3
 CO153765 490 bp mRNA linear EST 05-OCT-2004
 LOCUS EN02756.5prime Exelixis FlyTag MN08 Bluescript Drosophila
 DEFINITION melanogaster cDNA clone EN02756 5, mRNA sequence.
 ACCESSION CO153765
 VERSION CO153765.1 GI:48907766
 KEYWORDS EST
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 490)
 Nakanishi, M., Muzong, C., Peterson, E., Laufer, A., Leung, W., Platt, D.
 and Swimmer, C.
 Exelixis FlyTag EST Project MN08 Library
 Unpublished (2004)
 COMMENT Contact: Stapleton, M.
 BDCP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
 Plate: EN.27 row: E column: 8
 High quality sequence stop: 401.
 Location/Qualifiers
 1..490
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="EN02756"
 /cell_line="mbn2"
 /clone_1lb="Exelixis FlyTag MN08 Bluescript"
 /note="Vector: pBluescript; Site_1: NotI; Site_2: XhoI;
 oligodT primed from LPS induced mbn2 cell line."

ORIGIN
 Query Match 24.7%; Score 47; DB 7; Length 490;
 Best Local Similarity 37.7%; Pred. No. 0.00082;
 Matches 69; Conservative 42; Mismatches 70; Indels 2; Gaps 2;

QY 7 AUGUGAUCUUGCUGCGGAGCAAAUUGCAGUAUAAUUCGCAAGUAGUCUAU 66
 DB 107 ATGTGATCTTGCTTCTTATAC-AATTTTGAGAGGTTAATGAAGAGTAGTGCTATT 165
 QY 67 GUUGAUAUCCGUAUUAUAGGCUCCAAAGUGGUAUAGCAAGCCCUAUC 126
 DB 166 TTAAATTAATTAAGTTAACTTAAGTTAGTTCTGATGAGTGCCTAT-TGGCAGCCCCATA 224
 QY 127 AAUUAUCUAGAGAAACUGUCUAUUGUUUAGAAGUUAUGUAGUCUUAACAGAACAAU 186
 DB 225 ATATCCAGAGACACCCCTCTGCTCTTATATGATTAAGTTGTCATTAGAAATGAAGAAAT 284
 QY 187 ACC 189
 DB 285 AAC 287

RESULT 4
 CO337794 539 bp mRNA linear EST 05-OCT-2004
 LOCUS EN15317.5prime Exelixis FlyTag MN08 Bluescript Drosophila
 DEFINITION melanogaster cDNA clone EN15317 5, mRNA sequence.
 ACCESSION CO337794
 VERSION CO337794.1 GI:49398069
 KEYWORDS EST
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 539)

AUTHORS Nakanishi, M., Muzong, C., Peterson, E., Laufer, A., Leung, W., Platt, D.

and Swimer, C.

EXELIXIS FLYTAG EST Project MN08 Library

JOURNAL Unpublished (2004)

COMMENT Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu

Plate: EN.153 row: B column: 5

High quality sequence stop: 474.

Location/Qualifiers

1. 539

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="EN15317"

/cell_line="mbn2"

/note="Vector: pBluescript, Site_1: NotI; Site_2: XhoI; oligodt primed from LPS induced mbn2 cell line."

ORIGIN

Query Match 24.7% Score 47; DB 7; Length 539;

Best Local Similarity 37.7% Pred. No. 0.00083;

Matches 69; Conservative 42; Mismatches 70; Indels 2; Gaps 2;

7 AUGUGAUCUCUGCGGAGAAAUUGCAGUAUAAUUCGACAGUAGUCUAU 66

90 ATGATCTCTCTCTCTCTATAC -AATTTGAGAGGTTAATAAGAAAGTAGTCTATT 148

67 GUGGAUACACCCUACUUAUUGUUAUAGCUCUCAAUUGGUGAUAAGCCCUUUC 126

149 TTAATTAATTAAGTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 207

127 AAUUAUCUGAAGACUGUGCUUAUUGAUAUUGAUAUUGAUAUUGAUAU 186

208 ATATCCAGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 267

187 ACC 189

268 AAC 270

RESULT 5

CA248993/LOCUS

SCSBFL105E02.b FL1 Saccharum officinarum cDNA clone SCSBFL105E02

3' mRNA sequence.

CA248993 CA248993.1 GI:35331208

EST.

Saccharum officinarum

Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD

clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum

complex.

1 (bases 1 to 714)

Vector: A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: patricia@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

<http://www.bcccenter.fcav.unesp.br>

Plate: 105 row: E column: 02

Seq primer: SP6 Promoter Primer.

Location/Qualifiers

1. 714

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCSBFL105E02"

/lab_host="DH10B"

/clone_lib="FL1"

/note="Organ: Inflorescence at beginning of development (1cm-long); Vector: pSPori; Site_1: SalI; Site_2: NotI; An unidirectional cDNA library generated from Inflorescence at beginning of development (1cm-long). cDNA was prepared from polyA+ mRNA using Superscript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a Sephadex CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Query Match 19.4% Score 36.8; DB 6; Length 714;

Best Local Similarity 30.1% Pred. No. 1.5;

Matches 41; Conservative 33; Mismatches 62; Indels 0; Gaps 0;

41 GUUAAAUAUCGACAGUAGUCUAUUGUGAUAUACCCGUAUUAUUGUUAUAGCUC 100

532 GTCATTAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 473

101 CAAGUUGGUGAUAUAGCAGCCUAUUAUUGAUAUUGAUAUUGAUAUUGAUAU 160

472 AAAGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 413

161 UAGUAGUCUUAAC 176

412 GCCCTAGTTCTCTAC 397

RESULT 6

CW675417/LOCUS

OP_Ba0075D07.f OP_Ba Oryza punctata genomic clone OP_Ba0075D07

5' genomic survey sequence.

CW675417 CW675417.1 GI:55143431

GSS.

Oryza punctata

Oryza punctata

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoidae; Oryzae; Oryza.

1 (bases 1 to 826)

Sam Miguel, P., Westerman, R., Kim, H., Yu, Y., Wisesek, M., Yost, D.,

Stum, D., Rao, K., Luo, M., Jetty, R., Kudra, D., Muller, C.,

Hartfield, J., Soderlund, C., Wang, R., and Jackson, S.A.

OMP Project - Purdue University

Unpublished (2004)

Contact: Scott A. Jackson

Jackson Laboratory

Purdue University

915 W. State St., West Lafayette, IN 47907, USA

Tel: 7654863621

Fax: 7654867255

Email: sjackson@purdue.edu

Basecalling by Phred version 0.020425.c. This sequence was derived

from the raw sequence read by clipping with Jvarkit version 1.198.

Bases 36-861 of the raw sequence (length 1079) were retained after

clipping.

PCR Primers

FORWARD: TAA TAC GAC TCA CTA TAG GG

BACKWARD: CAC TCA TTA GGC ACC CCA

and adult (pooled)"
 /lab_host="DH10B"
 /clone_lib="normalized Keck-Tagu library SB02"
 /note="Organ: brain; Vector: pBS II SK(+); Site: 1:
 EcoRI(5' side of insert); Site 2: NotI (3' side of
 insert); The library was constructed and normalized as
 described by Bonaldo, M.F., Lennon, G. and Soares, M.B.
 (1996), Genome Research 6(9): 791-806. An identifying tag
 was added at the 3' during cDNA synthesis:
 InsertAAAAAAAAAAAAAAAAATGCA."

ORIGIN

Query Match 18.6%; Score 35.4; DB 7; Length 751;
 Best Local Similarity 33.5%; Pred. No. 4.3; Mismatches 29; Indels 86; Gaps 0;

Db 532 CCAGGAGCTGATCAAGATGAGATGATGCGCATCTCTTTCAGAAATTTGTAAGGTTT 591
 GY 1 CCACCAUUGUAGUCUUGCGGAGCAAAUUVUGACAGUAAUUAUCUGACAGUGU 60
 Db 592 GGAGAAAGGTGATGATGAGAGAAATGACATTCGATTCCTTCATGATGAGAGAG 651
 GY 61 GCUAUGUGAUAUACCGUACCUAUAUGUUAUGUUAUGUUAUGUUAUGUUAUGUUA 120
 Db 592 GGAGAAAGGTGATGATGAGAGAAATGACATTCGATTCCTTCATGATGAGAGAG 651
 GY 121 CCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 173
 Db 652 CCTATCTCTTCTACTCCACAGTGTATGAAATGAACTTCTTCTCTTA 704

RESULT 9

BX120183

DEFINITION Danio rerio genomic clone DKEX-69K2, genomic survey sequence.
 VERSION BX120183.1 GI:27951099
 KEYWORDS GSS.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio

LOCUS

BX120183

DEFINITION Danio rerio genomic clone DKEX-69K2, genomic survey sequence.
 VERSION BX120183.1 GI:27951099
 KEYWORDS GSS.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
 Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk Unpublished
 This sequence was generated from the T7 end of BAC 69K2. 69K2 is
 part of the Daniokey BAC library created by R. Plasterk and N.V.
 Keygene. Further details:
 http://www.sanger.ac.uk/projects/D_rerio/
 Location/Qualifiers

FEATURES

source

1. 608
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="DKEX-69K2"
 /issue_type="Teetle"
 /note="vector pindigobac-536"

ORIGIN

Query Match 18.3%; Score 34.8; DB 10; Length 608;
 Best Local Similarity 34.3%; Pred. No. 6.4; Mismatches 42; Indels 0; Gaps 0;

Db 3 AACAAUGUAGUCUUGCGGAGCAAAUUVUGACAGUAAUUAUCUGACAGUGUC 62
 GY 55 AACAAATGTTGATGATGAGAGAAATATTTGCTTCTATTAATTTAATTAATG 114
 Db 63 UAUGUGGAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 104
 GY 115 TTAGAGATGAATATTTTACATGTTAAATTAATTAATTAATTAATTAATTAAT 156

RESULT 10

BG499554

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BG499554 1228 bp mRNA linear EST 27-MAR-2001
 602546703F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4669033 5',
 mRNA sequence.
 BG499554
 BG499554.1 GI:13461071
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Homo
 1 (bases 1 to 1228)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/BTP
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LCM1480 row: e column: 02
 High quality sequence stop: 2.
 Location/Qualifiers

FEATURES

source

1. 1228
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4669033"
 /issue_type="adenocarcinoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_60"
 /note="Organ: prostate; Vector: pDNR-LIB (Clontech);
 Site 1: SfiI (ggcgccctggcc); Site 2: SfiI
 (ggccatcggcc); Double-stranded cDNA was prepared from
 cell line RNA. 5' and 3' adaptors were used in cloning as
 follows: 5' adaptor sequence: 5'-CAGCGCCATTATGGCC-3' and
 3' adaptor sequence:
 5'-ATTCTAAGCGCCGAGCGCGCCGACATG-dT(30)-BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.5
 kb (range 0.9-4.0 kb). 14/15 clones contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC library."

ORIGIN

Query Match 18.3%; Score 34.8; DB 2; Length 1228;
 Best Local Similarity 35.4%; Pred. No. 7.3; Mismatches 56; Conservative 25; Mismatches 77; Indels 0; Gaps 0;

GY 23 GGAGGCAAAUUVUGACAGUAAUUAUCUGAUAUGUCUUAUUAUCACCGUAC 82
 Db 484 GGTAAACAAGTCTATCGGTGTATGACATGAAGGGGTGCTCTAGGGAATCAACAAT 543
 GY 83 CUUUVUAGUUAUUAUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 142
 Db 544 CAATCTAGGTTGCTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 603
 GY 143 GUUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 180
 Db 604 GGGTAAAGTATAGGTGTAACAAGTAACTATCAAGAA 641

RESULT 11

CPA562935

LOCUS

DEFINITION

CPA562935 601 bp DNA linear GSS 11-AUG-2003
 Cryptosporidium parvum GSS, PAC clone pica_0012_e08, T7 end

Db	74	GACCTTCGTAACCTTCAAAATTATTGCGCATAGTCTATGTTTAAATAAGGTTT	15
Qy	169	CUCUAAACAGAAC	182
Db	14	ATGAAAAAAAAA	1

Search completed: December 22, 2005, 15:21:08
 Job time : 191.51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 07:55:17 ; Search time 79.9079 Seconds
(without alignment)
4226.575 Million cell updates/sec

Title: US-10-088-750C-6

Perfect score: 190
1 ccaacagaugauucuuu...cuaacagaacaaauuacuu 190

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/1.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCPUS.COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP.COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE.COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	24.7	1345	3	US-09-270-767-1072
2	47	24.7	1345	3	Sequence 1072, Ap
3	33	17.4	2701	3	Sequence 16354, A
4	32	16.8	87205	3	US-09-181-612-3
5	30.8	16.2	36661	3	US-09-949-016-13430
6	30	15.8	601	3	US-09-949-016-16102
7	30	15.8	250352	3	US-09-949-016-108154
8	29.6	15.6	241	3	US-09-949-016-14724
9	29.2	15.4	51589	3	US-09-004-113-17
10	29	15.3	42954	3	US-09-949-016-16936
11	29	15.3	42954	3	US-09-949-016-17123
12	28.8	15.2	353	3	US-09-270-767-11124
13	28.8	15.2	353	3	US-09-270-767-5518
14	28.8	15.2	353	3	US-09-270-767-20800
15	28.8	15.2	601	3	US-09-949-016-43832
16	28.8	15.2	601	3	US-09-949-016-43877
17	28.8	15.2	601	3	US-09-949-016-52449
18	28.8	15.2	601	3	US-09-949-016-52494
19	28.8	15.2	4209	3	US-09-248-796A-400
20	28.8	15.2	44100	3	US-09-719-554-63
21	28.8	15.2	45484	3	US-09-949-016-12967
22	28.8	15.2	45164	3	US-09-949-016-12985
23	28.8	15.2	45164	3	US-09-949-016-12986
24	28.8	15.2	49164	3	US-09-949-016-13262
					Sequence 13263, A

25	28.6	15.1	601	3	US-09-949-016-42431	Sequence 42431, A
26	28.6	15.1	1569	3	US-09-328-352-3981	Sequence 3281, Ap
27	28.6	15.1	49526	3	US-09-949-016-12959	Sequence 12959, A
28	28.6	15.1	108060	3	US-09-949-016-13159	Sequence 13159, A
29	28.6	15.1	121234	3	US-09-949-016-14142	Sequence 14142, A
30	28.6	15.1	140315	3	US-09-949-016-14141	Sequence 14141, A
31	28.6	15.1	147840	3	US-09-949-016-15236	Sequence 15236, A
32	28.6	15.1	265038	3	US-09-949-016-15719	Sequence 15719, A
33	28.4	14.9	783	3	US-09-302-769-15	Sequence 15, Appl
34	28.4	14.9	783	3	US-08-962-560C-15	Sequence 15, Appl
35	28.4	14.9	27663	3	US-09-949-016-16160	Sequence 16160, A
36	28.4	14.9	39318	3	US-09-949-016-13798	Sequence 13798, A
37	28.4	14.9	44499	3	US-09-949-016-12711	Sequence 12711, A
38	28.4	14.9	70014	3	US-09-949-016-17110	Sequence 17110, A
39	28.2	14.8	2056	3	US-08-158-735A-1	Sequence 1, Appl
40	28.2	14.8	2070	3	US-09-382-256-13	Sequence 13, Appl
41	28.2	14.8	2070	3	US-09-382-256-13	Sequence 13, Appl
42	28.2	14.8	2070	3	US-08-436-265-13	Sequence 13, Appl
43	28.2	14.8	2070	3	US-09-679-187-13	Sequence 13, Appl
44	28.2	14.8	2070	3	US-09-267-963D-13	Sequence 13, Appl
45	28.2	14.8	2402	3	US-08-462-467B-11	Sequence 11, Appl

ALIGNMENTS

```
RESULT 1
US-09-270-767-1072
Sequence 1072, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1072
; LENGTH: 1345
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1072

Query Match      24.7%; Score 47; DB 3; Length 1345;
Best Local Similarity 37.7%; Pred. No. 1.4e-05;
Matches 69; Conservative 42; Mismatches 70; Indels 2; Gaps 2;

QY 7 AUGUGAUCUUGCGGAGCAAAUUGGACAGUAAAUUUGGACAGUAGUGCUU 66
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 523 ATGTGATCTTGCTCTTATAC-AAITTTGAGAGTTAATAAGAAAGTAGTCTATT 581
   ||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 67 GUGGAAUACCGUACCUUUGAGUUUAGCCUCCAGAGUGGAGUAGCAGCCCUU 126
   ||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 582 TTATATATTAGTTAATTAATTAGTTTACGTTCAGATCCAT-TCGAGAGCCCA 640
   ||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 127 AAUUCUAGAGAAUUGUGUUAUUGAGAUAGUAGUUCUUAACCAAUU 186
   ||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 641 ATATCCAGACACCTCTGCTCTTATGATTAATTAATTAATTAATTAAT 700
   ||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 187 ACC 189
   ||
DB 701 AAC 703

RESULT 2
US-09-270-767-16354
Sequence 16354, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
```



```

: CURRENT APPLICATION NUMBER: US/09/270,767
: CURRENT FILING DATE: 1999-03-17
: NUMBER OF SEQ ID NOS: 62517
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 16354
: LENGTH: 1345
: TYPE: DNA
: ORGANISM: Drosophila melanogaster
US-09-270-767-16354

Query Match      24.7%; Score 47; DB 3; Length 1345;
Best Local Similarity 37.7%; Pred. No. 1.4e-05;
Matches 69; Conservative 42; Mismatches 70; Indels 2; Gaps 2

QY 7 AUGGAAUCUUGUCUGCGGAGGCAAAUUGCAGAUAAAUCGCAAGUUGUCUAAU 66
Db 523 ATGCGATCTTCCTCTCCTTATAC-AAATTGAGAGGTTAATAGAAAGAAAGTGTCTATT 581
QY 67 GUGGAAUACCCGACCUAAUUAAGUUUACGUCCAAGUCCGUGAUAAGACCCUAC 126
Db 582 TTAATTAATTAGGTTAACTATTAGTTTAACTGTTCAGATGCGCTAT-TGGCAGCCCATTA 640
QY 127 AAUUAUAGAGGAACUGUGCUAUUGUUGAAGAUUAGUAUUCUUAACGAACAAU 186
Db 641 ATATCCAGAGACACCTCTCTGCTCTTATATGATTAAGTTGTCATTGTAATTAAGAAAT 700
QY 187 ACC 189
Db 701 AAC 703

RESULT 3
US-10-181-612-3/c
: Sequence 3, Application US/10181612
: Patent No. 6962988
: GENERAL INFORMATION:
: APPLICANT: Coomrod, Scott A
: APPLICANT: Herr, John C
: APPLICANT: Wright, Paul
: TITLE OF INVENTION: Egg Specific Surface Proteins
: FILE REFERENCE: 00498-07
: CURRENT APPLICATION NUMBER: US/10/181,612
: CURRENT FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: US 60/177,123
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 3
: LENGTH: 2701
: TYPE: DNA
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (2337)..(2337)
: OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (2340)..(2340)
: OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (2366)..(2366)
: OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (2382)..(2382)
: OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (2447)..(2447)
: OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
: NAME/KEY: misc_feature

```

```

LOCATION: (2453)..(2453)
OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2481)..(2481)
OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2509)..(2509)
OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2525)..(2525)
OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2556)..(2556)
OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2577)..(2577)
OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2580)..(2580)
OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2630)..(2630)
OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2666)..(2666)
OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
US-10-181-612-3
Query Match      17.4%; Score 33; DB 3; Length 2701;
Best Local Similarity 32.4%; Pred. No. 0.67;
Matches 61; Conservative 29; Mismatches 98; Indels 0; Gaps 0;
QY      3 AACAAUGUAGUCUGUCUGUCGAGCGAGCAAAUUVUGACAGUUAUAUUCGACAGUAGUC 62
Db      2376 AAAAAAATTATNTAAAGCTGCGGTATTAATGAGCANTTAAAGATGAGAAAAATTTA 2317
QY      63 UAUUGUGAUAUCACCGGACCUAAUUVUAGUUVUACGUCGCAAGUCGUGAUGACAGCCC 122
Db      2316 TGAGCGTTGGCTCTCAATCATCATCTATTCATTTTAAACCAATGATGTGGACAGGTGGC 2257
QY      123 UATCAUAUUCUGAAGAAUCUGUCGUAUUVUUAAGAAGUATAGUAGUUGUCUUAACAGAA 182
Db      2256 TTTCATGATCAAGAGAGTTGTGCTTAGCTTTAGGTGCTCCCTGAGGTATCTGAAGACATA 2197
QY      183 AUUUAACU 190
Db      2196 CTTTGCT 2189
RESULT 4
US-09-949-016-13430/c
; Sequence 13430, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```


QY 83 CUATUUGAGUUUAGCGUCCAGAGUCCGUGAGUAGACGCCUUAUUCUAGAGAACTU 142
DB 47684 GACTTGATTGGTACGACGAAAAGAGATTTCACATCACTCACTGATGAAAA 47743
QY 143 GUGCUGUUGUAGAAGUUVUAGUGUCUCUAAACAGAA 180
DB 47744 ATTTATCTTATTAACAGCATATGTAATTTAATAGGA 47781

RESULT 8
US-09-004-113-17
Sequence 17, Application US/0904113
Patent No. 6028185
GENERAL INFORMATION:
APPLICANT: Ozias-Akins, Peggy
APPLICANT: Hanna, Wayne W.
APPLICANT: Roche, Dominique
TITLE OF INVENTION: Nucleic Acid Markers for
TITLE OF INVENTION: Apospory-Specific Genomic Region
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Gall E. Poulos
STREET: Room 407, Bldg. 005, BARC-W
CITY: Beltsville
STATE: Maryland
COUNTRY: USA
ZIP: 20705
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/004,113
APPLICATION NUMBER: US/09/004,113
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Poulos, Gall E.
REGISTRATION NUMBER: 36,327
REFERENCE/DOCKET NUMBER: 0008, 98
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-504-5302
TELEFAX: 301-504-5060
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 241 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pennisetum squamulatum
IMMEDIATE SOURCE:
CLONE: M10 3' end
US-09-004-113-17

Query Match 15.6%, Score 29.6; DB 3; Length 241;
Best Local Similarity 35.2%; Pred. No. 4.2;
Matches 38; Conservative 21; Mismatches 49; Indels 0; Gaps 0;
QY 79 GUACCUUUUUGUUUAGCGUCCAGAGUCCGUGAGUAGACGCCUUAUUCUAGAG 138
DB 14 GTACCAAAATTATGTTGACAAATTCATTTCTAGAAATTTTGAAGCACATTAATTATTATG 73
QY 139 AACUGUCUAGUUGUAGAAGUUVUAGUGUCUCUAAACAGAACAUUVU 186
DB 74 AATATATATATCATGAGCATGAGATGTAAATTCGATCATATAAGAAATTT 121

RESULT 9

US-09-949-016-16936/c
Sequence 16936, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16936
LENGTH: 51589
TYPE: DNA
ORGANISM: Human
US-09-949-016-16936

Query Match 15.4%, Score 29.2; DB 3; Length 51589;
Best Local Similarity 41.5%; Pred. No. 29;
Matches 34; Conservative 15; Mismatches 33; Indels 0; Gaps 0;

QY 43 AUAUAAUUCUGCAAGUAGUCUATUUGUAGAUACCGUACCUUUVUAGUUGUUGCGUCC 102
DB 13043 AAAGAAATCAGCTGGGATCTATTTGAAAGAAAGAAAGATAGATGTGGGTCTAGCGTCCA 12984
QY 103 AGAUCGUGUAGUAGACGCCCUA 124
DB 12983 ATATTGAGTTAAGAGACTTGTA 12962

RESULT 10
US-09-949-016-17123
Sequence 17123, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17123
LENGTH: 42954
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(42954)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17123

Query Match 15.3%, Score 29; DB 3; Length 42954;
Best Local Similarity 39.6%; Pred. No. 32;
Matches 40; Conservative 16; Mismatches 45; Indels 0; Gaps 0;

QY 42 UAUAUAAUUCUGCAAGUAGUCUATUUGUAGAUACCGUACCUUUVUAGUUGUUGCGUCC 101
DB 30513 TAAAAATTAAGCAGCGCGGTGTATGTCTGTAGTCCAGCTACTGTGGGGTGTAGGCGAG 30572

Qy 102 AAGATCGGUGGAUAGCAGCCCUAUCAUAUUCUAGAGAACTU 142
 |||: |||: |||:
 Db 30573 AAGATCACTTGAGCCCAAGCGTTTAAAGCTCAACTCAAACT 30613

RESULT 11

```

US-09-949-016-17124
: Sequence 17124, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001107
: CURRENT APPLICATION NUMBER: US/09/949, 016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 17124
: LENGTH: 42954
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(42954)
: OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17124

```

Query Match	15.3%	Score 29;	DB 3;	Length 42954;
Best Local Similarity	39.6%	Pred. No. 32;		
Matches 40;	Conservative 16;	Mismatches 45;	Indels	

42 UUUUAAAUUUCUGCAAGUGUGCUUUUUUGUGAAUACAACGUACCUUUUUUAGUGUUUACCCUCC 101
 30513 TAAAAAATTAGCCAGGCGCTGGTATGCTCTGTAGTCCACACTATCGGGGTGGTTGAGCCAGG 30572

Qy 102 AAGAUCCGUGGAUAGCAGCCCUAUCAUAUUCUAGAGAACT 142
 ||||:|:||||:|:|:|:
 Db 30573 AAGATCACTTGAGCCAGGCGTTTAGGCTCAACTCAAACT 30613

RESULT 12

```

US-09-270-767-5518/c
; Sequence 5518, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-034
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5518
; LENGTH: 353
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-5518

```

Query Match	15.2%	Score 28.8	DB 3	Length 353
Best Local Similarity	38.9%	Pred. No. 8.7		
Matches 28, Conservative	17	Mismatches 27	Indels 0	Gaps 0

Oy 30 AAATUUGCAGUAUAAAUUCUGCAAGUAGUCUAUTUGUGAUAUCACCGUACCUAUUA 89
| : : | : : | : : | : : | : : | : : | : : | : :
Db 333 ATATTTCCTTCCGTATAACTCATCTGAGACTAAATGATTGTTAGTTGTACATATATT 274

Qy	90	GGUUACGCUC	101
		: :	
Db	273	TTTATAGGACC	262

RESULT 13

```

US-09-270-767-20800/c
; Sequence 20800, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20800
; LENGTH: 353
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-20800

```

Query Match	15.24;	Score 28.8;	DB 3;	Length 353;
Best Local Similarity	38.94;	Pred. NO. 8.7;		
Matches 28;	Conservative 17;	Mismatches 27;	Indels 0;	Gaps 0;

Oy 30 AAAUUUGCACAGUAAAAAUCUGCAGAUAUGUCUAUUGUUGGAUUCACCGUACCUAUUUA 89
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 333 ATATTGGCTTCGTATAACTCATCATGTCGAGCTAAATGATTGATTGTCATATATT 274

Qy	90	GGUUACGCUC	101
		∶ ∶ ∣ ∣ ∣	
Db	273	TTTATAGGAGCC	262

RESULT 14

```

US-09-949-016-43832/C
: Sequence 43832, Application US/09943016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMERISMS IN KNOWN GENES ASSOCIATED
: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949.016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 43832
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-43832

```

Query Match	15.2%	Score 28.8	DB 3	Length 601
Best Local Similarity	34.0%	Pred. No. 10		
Matches 49; Conservative	23	Mismatches	72	Indels 0
				Gaps 0

QY 14 CUCUCUUGGGAGGCAAAAVUUGCACAAGUAAAAUUCGCAAGUAGUGCUAUUGUGGAA 73
| : : | | | | | | | | : : : : : |
DB 574 CACGTTTATTAGGCCAAACCGTGACGAGCAGCTAGATTATTACTTTTAACTGTAA 515

```
Oy      74 UCACCGUAGCUAUAUGUUUAGCCGCCAAGAUCGUGAUAGCAGCCUAUCAUAUUCU   133
       ||::||::|:||::||::||::||::||::||::||::||::||::||::||:
Db     514 GTACTGTACATAGTTAAATTCTTCACTTAAGTGTCGTGAACCACAACCTGGAACTACT   455
```


QY 134 AGGAGAACUGUCUAUUGUUUGAA 157
Db 454 ATTCATGCGAGGCAATGATGATTA 431

RESULT 15

US-09-949-016-43877/C
; Sequence 43877, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43877
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-43877

Query Match 15.2%; Score 28.8; DB 3; Length 601;

Best Local Similarity 34.0%; Pred. No. 10; Mismatches 72; Indels 0; Gaps 0;

QY 14 CUGGCUUGCGGAGGCAAAUUVUGCAAGUUAUAAUUCGACAGUAGUGCUAUUGUGGAA 73
Db 574 CATGTTTATTATTAGCGCAACCGGTACGAGCAGCTAGATTATTTACTTTACTGTTAA 515
QY 74 UCACCGUACCUAUUUGAGUUUAGCGUCCAGAGAUUGUGAGUAGCAGCCCUAUCAUUAUCU 133
Db 514 GTAAGTACATAGTAAATTTATTCATCTAAGTGTGTGTGAGCCAACTGTGAATTACT 455
QY 134 AGGAGAACUGUCUAUUGUUUGAA 157
Db 454 ATTCATGCGAGGCAATGATGATTA 431

Search completed: December 22, 2005, 15:30:27
Job time : 82.9079 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 08:09:01 / Search time 442.264 Seconds
(Without alignments)
3552.588 Million cell updates/sec

Title: US-10-088-750C-6
Perfect score: 190
Sequence: 1 ccacacgaugauccuugcu.....cuaacagacaaauuaccu 190

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA Main:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	190	100.0	202	7	US-10-088-750B-6 Sequence 6, Appli
2	63.4	33.4	200	7	US-10-088-750B-4 Sequence 4, Appli
3	45.4	23.9	201	7	US-10-088-750B-3 Sequence 3, Appli
4	37	19.5	199	7	US-10-088-750B-2 Sequence 2, Appli
5	35.6	18.7	653	4	US-09-925-065A-611338 Sequence 611338,
6	35.6	18.2	459	4	US-09-925-065A-151690 Sequence 151690,
7	34.6	18.0	640	5	US-10-027-632-219841 Sequence 219841,
8	34.2	18.0	640	6	US-10-027-632-219841 Sequence 219841,
9	34.2	18.0	640	6	US-10-027-632-219841 Sequence 219841,
10	33	17.4	2701	6	US-10-181-612-3 Sequence 3, Appli
11	33	16.8	95832	8	US-10-741-600-17662 Sequence 17662, A
12	31.6	16.6	1626	4	US-09-925-065A-91880 Sequence 91880, A
13	31.6	16.6	1626	4	US-09-925-065A-91880 Sequence 91880, A
14	31.6	16.6	1626	4	US-09-925-065A-91880 Sequence 91880, A
15	31.2	16.4	2000	3	US-09-938-842A-4112 Sequence 4112, Ap
16	31.2	16.4	2000	3	US-09-938-842A-4112 Sequence 4112, Ap
17	31	16.3	1140	7	US-10-424-599-45821 Sequence 45821, A
18	31	16.3	50353	8	US-10-741-600-17699 Sequence 17699, A
19	31	16.3	380963	9	US-10-737-082-5 Sequence 5, Appli
20	31	16.3	380963	9	US-10-737-082-5 Sequence 5, Appli
21	30.8	16.2	421	7	US-10-424-599-105537 Sequence 105537,
22	30.8	16.2	52661	3	US-09-997-722-142 Sequence 142, App
23	30.8	16.2	74822	7	US-10-240-425-1121 Sequence 1121, Ap

24	30.8	16.2	74822	9	US-10-756-149-1042 Sequence 1042, Ap
25	30.8	16.2	91507	10	US-11-097-143-15691 Sequence 15691, A
26	30.4	16.0	573	4	US-09-925-065A-822287 Sequence 822287,
27	30.4	16.0	573	4	US-09-925-065A-822288 Sequence 822288,
28	30.4	16.0	600	9	US-10-972-079-67201 Sequence 67201, A
29	30.2	15.9	1190	8	US-10-767-795-4793 Sequence 4793, Ap
30	30.2	15.9	2096	7	US-10-437-963-53757 Sequence 53757, A
31	30.2	15.9	38258	6	US-10-365-319-1 Sequence 1, Appli
32	30.2	15.9	39005	8	US-10-753-901-6 Sequence 6, Appli
33	30.2	15.9	39005	8	US-10-753-901-6 Sequence 6, Appli
34	30	15.8	438	4	US-09-925-065A-73912 Sequence 73912, A
35	30	15.8	557	4	US-09-925-065A-573584 Sequence 573584,
36	29.8	15.7	407	8	US-10-674-124A-8199 Sequence 8199, Ap
37	29.8	15.7	557	5	US-10-027-632-257350 Sequence 257350,
38	29.8	15.7	557	5	US-10-027-632-257351 Sequence 257351,
39	29.8	15.7	557	6	US-10-027-632-257350 Sequence 257350,
40	29.8	15.7	557	6	US-10-027-632-257351 Sequence 257351,
41	29.8	15.7	1093	7	US-10-424-599-70334 Sequence 70334, A
42	29.8	15.7	3036	5	US-10-027-632-112419 Sequence 112419,
43	29.8	15.7	3036	5	US-10-027-632-112420 Sequence 112420,
44	29.8	15.7	3036	5	US-10-027-632-112421 Sequence 112421,
45	29.8	15.7	3036	6	US-10-027-632-112419 Sequence 112419,

ALIGNMENTS

RESULT 1
US-10-088-750B-6
Sequence 6, Application US/10088750B
Publication No. US20040166486A1
GENERAL INFORMATION:
APPLICANT: NAKASHIMA, Nobuhiko
TITLE OF INVENTION: A Novel Higher-Order Structure With Promoting Translation
FILE REFERENCE: 3190-015
CURRENT APPLICATION NUMBER: US/10/088, 750B
PRIOR FILING DATE: 2002-03-20
PRIOR APPLICATION NUMBER: JP P2001-016746
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: PCT/JP01/00641
PRIOR FILING DATE: 2001-01-31
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.2
SEQ ID NO 6
LENGTH: 202
TYPE: RNA
ORGANISM: Black Queen-Cell Virus
US-10-088-750B-6
Query Match 100.0%; Score 190; DB 7; Length 202;
Best Local Similarity 100.0%; Pred. No. 2.1e-52;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CCAACAUGAUGAUCUGCGGAGCAAAUUTUGCAGAUAAUUCGACAGUAGU 60
DB 1 CCAACAUGAUGAUCUGCGGAGCAAAUUTUGCAGAUAAUUCGACAGUAGU 60
OY 61 GCUAUGUGGAAUACCGUACCUAUVUAGUUVUACGUCGACAGUGGUGAUGAGC 120
DB 61 GCUAUGUGGAAUACCGUACCUAUVUAGUUVUACGUCGACAGUGGUGAUGAGC 120
OY 121 CCUACAUUUCUAGGAGAACTGUGCUAUVUUTUAGAAUAGUAGUACCUAAACGAA 180
DB 121 CCUACAUUUCUAGGAGAACTGUGCUAUVUUTUAGAAUAGUAGUACCUAAACGAA 180
OY 181 CAAUUAUCCU 190
DB 181 CAAUUAUCCU 190
RESULT 2


```

? PRIOR FILING DATE: 1999-11-23
? PRIOR APPLICATION NUMBER: US 60/156,358
? PRIOR FILING DATE: 1999-09-28
? PRIOR APPLICATION NUMBER: US 60/146,002
? PRIOR FILING DATE: 1999-08-09
? NUMBER OF SEQ ID NOS: 325720
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO: 219841
? LENGTH: 640
? TYPE: DNA
? ORGANISM: Human
? US-10-027-632-219841

```

Query Match	18.0%;	Score 34.2;	DB 5;	Length 640;
Best Local Similarity	29.4%;	Pred. No. 1.3;		
Matches	42;	Conservative 33;	Mismatches 68;	Indels 0; Gaps 0

Oy	42	U A U A A A A C U G C A A U G U G C A U U U G U G A U C A C C C G A C C A U U U A G U G U N A G C U C C	101
	:	: : : : : :	:
	:	: : : : : :	:
Db	92	T A T A G T T G T A C A T G C T G A T G T A T A T T C A T A C T T G T C A T A A G T T T T T C A C T C T T A	151
Oy	102	A A G A U C G U G G A U D A G C A G C C C A U C A A U U C U G A G A A C U G U G C U A U G U U A G A A G A U	161
	:	: : : : : :	:
	:	: : : : : :	:
Db	152	A A A A T G T A T A G T A G T A G T T G C A C A T A T T T G S G G G C A T A T G C A T T T T A T A C C T T T	211
Oy	162	A G U N G U C U C U A A C A G A C A A U	184
	:	: : : : : :	:
	:	: : : : : :	:
Db	212	A T A T A G T G T A T A G T A T A A A A T T	234

```

RESULT 9
US-10-027-632-219841
; Sequence 219841, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 219841
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-219841

```

	Query Match	18.0%;	Score 34.2;	DB 6;	Length 640;
	Best Local Similarity	29.4%;	Pred. No. 1.3;		
	Matches	42;	Conservative	33;	Mismatches 68; Indels 0; Gaps 0;
Oy	42	UAAUAACUCGCAAGUAGUGCUAUGUGGAUACCCGUACUUUUUAGUUUAGCUGC	101		
	:	: : : : :	:	:	: : : : :
Dd	92	TATGGTTGTGACAGTGCCTGATGTATTTCATATCTTGCTCAATGAGTTTTTTCACCTTCA	151		
OY	102	AAGAUCGCGAUAGACACCUCUAUCAUUCUAGGAGACUGGCUCAUGUUUAGAAGAU	161		
	:	: : : : :	:	:	: : : : :

Db	152	AAATGATATGATGATGATGATATTTGGGGGCATATGTCATTTTGGATACCTT	211
Qy	162	AGGUAUCUCUAAACAGAACAAU	184
Db	212	ATATAGTGTATATGATATAAATT	234

```

RESULT 10
US-10-181-612-3/C
; Sequence 3, Application US/10181612
; Publication No. US20030186369A1
; GENERAL INFORMATION:
; APPLICANT: Coonrod, Scott A
; APPLICANT: Herr, John C
; APPLICANT: Wright, Paul
; TITLE OF INVENTION: Egg Specific Surface Proteins
; FILE REFERENCE: 00498-07
; CURRENT APPLICATION NUMBER: US/10/181,612
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/177,123
; PRIOR FILING DATE: 2000-01-20
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2701
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2337)..(2337)
; OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2340)..(2340)
; OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2366)..(2366)
; OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2382)..(2382)
; OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2447)..(2447)
; OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2453)..(2453)
; OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2481)..(2481)
; OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2509)..(2509)
; OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2525)..(2525)
; OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2556)..(2556)
; OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (2577)..(2577)
; OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
; FEATURE:
; NAME/KEY: misc.feature

```



```
LOCATION: (2580)..(2580)
OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
FEATURE:
NAME/KEY: misc feature
LOCATION: (2630)..(2630)
OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
FEATURE:
NAME/KEY: misc feature
LOCATION: (2666)..(2666)
OTHER INFORMATION: "n" meaning any nucleotide selected from A, T, G or C
US-10-181-612-3

Query Match
Best Local Similarity 17.4%; Score 33; DB 6; Length 2701;
Matches 61; Conservative 29; Mismatches 98; Indels 0; Gaps 0;

OY 3 AACAAUGGAGUCCUGGCGGAGGCAAAUUTUGCAGUAAAUAUUCGCAAGUAGUGC 62
Db 2376 AAAAAATTATTAAGACTGGGTGTATAAATGACGANTNAAGATGAGAGAAAAATTTA 2317

OY 63 UAUGUGGAAUACCGGACCUAUUUAGUUUAGCCUCCAGAGUGGUGAGACGCC 122
Db 2316 TGAGGCTTGGCTCAATCACTACATTTTAACACATGATGTTGGACAGGGTGC 2257

OY 123 UAUCAAUUCUAGGAGAACUGUGCUAUGUUUAGAUAUAGUAGUCUCUAAACAGACA 182
Db 2356 TTTGATGATCAAGAGGTTGTCTGCTTTAGGGTCCCTGAGGATCTGAAGACATA 2197

OY 183 AUUUAACU 190
Db 2196 CTTTGCCCT 2189

RESULT 11
US-10-741-600-17662/c
Sequence 17662, Application US/10741600
Publication No. US20050026169A1
GENERAL INFORMATION:
APPLICANT: CARCILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CLO001499
CURRENT APPLICATION NUMBER: US/10/741,600
CURRENT FILING DATE: 2003-12-22
NUMBER OF SEQ ID NOS: 73997
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17662
LENGTH: 95832
TYPE: DNA
ORGANISM: Homo sapiens
US-10-741-600-17662

Query Match
Best Local Similarity 16.8%; Score 32; DB 8; Length 95832;
Matches 47; Conservative 33; Mismatches 80; Indels 0; Gaps 0;

OY 3 AACAAUGGAGUCCUGGCGGAGGCAAAUUTUGCAGUAAAUAUUCGCAAGUAGUGC 62
Db 41305 ATCGAATGATTTTATAGAGAAATTCCTTCATATATTTTCAGAGCTGGAACCTTA 41146

OY 63 UAUGUGGAAUACCGGACCUAUUUAGUUUAGCCUCCAGAGUGGUGAGACGCC 122
Db 41145 TTTAGTTGAAATTTTCCTGCTCAAGACAGTTTAACATCATATATATATCATCATCA 41086

OY 123 UAUCAAUUCUAGGAGAACUGUGCUAUGUUUAGAUAUAGUUA 162
Db 41085 TTTTATCTCCATACAAATTAAGTATTTTCAAAATTTTA 41046

RESULT 12
US-09-925-065A-91879
Sequence 91879, Application US/09925065A
Publication No. US20050228172A9

US-10-088-750c-6.rnbp
```

```
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 91879
LENGTH: 1626
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-91879

Query Match
Best Local Similarity 16.6%; Score 31.6; DB 4; Length 1626;
Matches 24; Conservative 16; Mismatches 14; Indels 0; Gaps 0;

OY 16 UGCUUGGAGGCAAAUUTUGCAGUAAAUAUUCGCAAGUAGUGCUAUGUGU 69
Db 1460 TGTGTGGAGGTAAATATACAGATATACATTTGCCAGTTTCTATTTT 1513

RESULT 13
US-09-925-065A-91880
Sequence 91880, Application US/09925065A
Publication No. US20050228172A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 91880
LENGTH: 1626
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-91880

Query Match
Best Local Similarity 16.6%; Score 31.6; DB 4; Length 1626;
Matches 24; Conservative 16; Mismatches 14; Indels 0; Gaps 0;

OY 16 UGCUUGGAGGCAAAUUTUGCAGUAAAUAUUCGCAAGUAGUGCUAUGUGU 69
Db 1460 TGTGTGGAGGTAAATATACAGATATACATTTGCCAGTTTCTATTTT 1513

RESULT 14
```



```
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 13305
LENGTH: 50353
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-13305

Query Match
Best Local Similarity 16.3%; Score 31; DB 6; Length 50353;
Matches 39; Conservative 22; Mismatches 50; Indels 0; Gaps 0;

Qy 58 AGUCUUAUUGGUAUACCCUACCUUAUUGUUAUAGCCUACAAGUCCGUGAGUAGC 117
Db 37707 AGGAGTCTTGATTAATTAATTAACCTTCTGAGCTTGTTCACACTGGAAGAAAGC 37648
Qy 118 AGCCUACAUAUUCAGAGAAUCUGUCUAUUGUUAUAGAAUUAUAGUAGU 168
Db 37647 AAGCTATGATTAATATCATACCGGGTTACATGTGAGCAATTAAGTAAAT 37597

RESULT 3
US-11-117-187-201/c
Sequence 201, Application US/11117187
Publication No. US20050266560A1
GENERAL INFORMATION:
APPLICANT: PREUSS, DAPHNE
TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
FILE REFERENCE: ARCD:109US
CURRENT APPLICATION NUMBER: US/11/117,187
CURRENT FILING DATE: 2005-04-28
PRIOR APPLICATION NUMBER: US/09/531,120
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/125,219
PRIOR FILING DATE: 1999-03-18
NUMBER OF SEQ ID NOS: 212
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 201
LENGTH: 80450
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-11-117-187-201

Query Match
Best Local Similarity 15.5%; Score 29.4; DB 7; Length 80450;
Matches 34; Conservative 20; Mismatches 41; Indels 0; Gaps 0;

Qy 24 GAGGCAAAUUVGACAGUAAUAAUUCGCAAGUAGUCUUAUUGGAAUACCCUAGC 83
Db 56544 GAAAGAAUAAUATTTCAATATATATTTCTTCAAGTATCTATAGAAATGTTAC 56485
Qy 84 UAUUAGUUAUACGUCUCCAGAUUCGUGAUUAGCA 118
Db 56484 TACTTTGTCTACTCTCTGCGTATGGAATACCA 56450

RESULT 4
US-10-750-185-21629
Sequence 21629, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: PANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
```

```
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 21629
LENGTH: 600
TYPE: DNA
ORGANISM: Bovine
US-10-750-185-21629

Query Match
Best Local Similarity 15.2%; Score 28.8; DB 6; Length 600;
Matches 33; Conservative 21; Mismatches 42; Indels 0; Gaps 0;

Qy 5 CAUUGAUUCUUCGCGAGGCAAAUUVGCAAGUAAUUAUUCGAAUAGUAGCUA 64
Db 88 CAACATTAATCTGTTGATTAATTAAGAAAGTTTAAATATTAATAAACTGTTAAAGTCAT 147
Qy 65 UUGUGAAUACCCGUAUUAUUGUUAUAGCCUC 100
Db 148 TTTTCTGTCAGATTACCTGTCTATGTTCCATTTC 183
```

```
RESULT 5
US-10-750-185-60944
Sequence 60944, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: PANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 60944
LENGTH: 1209
TYPE: DNA
ORGANISM: Bovine
US-10-750-185-60944

Query Match
Best Local Similarity 15.2%; Score 28.8; DB 6; Length 1209;
Matches 33; Conservative 21; Mismatches 42; Indels 0; Gaps 0;

Qy 5 CAUUGAUUCUUCGCGAGGCAAAUUVGCAAGUAAUUAUUCGAAUAGUAGCUA 64
Db 370 CAACATTAATCTGTTGATTAATTAAGAAAGTTTAAATATTAATAAACTGTTAAAGTCAT 429
Qy 65 UUGUGAAUACCCGUAUUAUUGUUAUAGCCUC 100
Db 430 TTTTCTGTCAGATTACCTGTCTATGTTCCATTTC 465
```

```
RESULT 6
US-10-750-185-45634
Sequence 45634, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
```



```

; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45634
; LENGTH: 2874
; TYPE: DNA
; ORGANISM: Bovine 19866881349204
US-10-750-185-45634

```

```

Query Match 15.2%; Score 28.8; DB 6; Length 2874;
Best Local Similarity 38.3%; Pred. No. 3.9;
Matches 46; Conservative 17; Mismatches 57; Indels 0; Gaps 0;

QY 1 CCAACAUGGAGUCUUGCGGAGGCAAAUUVUGCAGUUAUAAUUCUGCAAGUAGU 60
DB 1797 CAATGATCTCAGCTCAGATATGCTAAAGATTCCCAAGTACACACGCAAGAAATA 1856
QY 61 GCUAUGUGGAAUACCGUACCUAUVUAGUUAUCCUGCAAGUGGAGUAGACAC 120
DB 1857 GCCATTAAAGGTCTAAGAAATTTGTTAACTTAAACCAAAATCAATAGACGACGC 1916

```

```

RESULT 7
US-10-995-561-13263/c
; Sequence 13263, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13263
; LENGTH: 57198
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13263

```

```

Query Match 15.2%; Score 28.8; DB 6; Length 57198;
Best Local Similarity 34.0%; Pred. No. 14;
Matches 49; Conservative 23; Mismatches 72; Indels 0; Gaps 0;

QY 14 CUUCGUCGCGAGCAAAUUVUGCAGUUAUAAUUCUGCAAGUAGUAGUAGUAGU 73
DB 28633 CATGTTATTTTAGGCAAAACCGGTACGAGCGAGTAAATTAATTAATTAATTA 28774
QY 74 UCACCGUACCUAUVUAGUUAUCCUGCAAGUGGAGUAGUAGUAGUAGUAGUAGU 133
DB 28773 GACTGTACATAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 28714
QY 134 AGAGAAACUGUGCUAUVUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 157
DB 28713 ATTCATGACGAGGCAATGATGATTA 28690

```

```

RESULT 8
US-10-995-561-13277/c
; Sequence 13277, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:

```

```

; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13277
; LENGTH: 173995
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13277

```

```

Query Match 15.2%; Score 28.8; DB 6; Length 173995;
Best Local Similarity 34.0%; Pred. No. 21;
Matches 49; Conservative 23; Mismatches 72; Indels 0; Gaps 0;

QY 14 CUUCGUCGCGAGCAAAUUVUGCAGUUAUAAUUCUGCAAGUAGUAGUAGUAGU 73
DB 28749 CATGTTATTTTAGGCAAAACCGGTACGAGCGAGTAAATTAATTAATTAATTA 28690
QY 74 UCACCGUACCUAUVUAGUUAUCCUGCAAGUGGAGUAGUAGUAGUAGUAGUAGU 133
DB 28689 GACTGTACATAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 28630
QY 134 AGAGAAACUGUGCUAUVUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 157
DB 28629 ATTCATGACGAGGCAATGATGATTA 28606

```

```

RESULT 9
US-10-750-185-40632
; Sequence 40632, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40632
; LENGTH: 952
; TYPE: DNA
; ORGANISM: Bovine 19866880791528
US-10-750-185-40632

```

```

Query Match 14.9%; Score 28.4; DB 6; Length 952;
Best Local Similarity 44.7%; Pred. No. 3.4;
Matches 42; Conservative 11; Mismatches 41; Indels 0; Gaps 0;

QY 35 UGCAAGUUAUAAUUCUGCAAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 94
DB 768 TCACGGTAAAGATTCGCTGACGTGAGAGATCCAGGTTCCGTTCTTCTGTTGGGA 827
QY 95 ACGUCCAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 128
DB 828 AATCCCTCGAAAGAGATGCAATCCACTCTA 861

```

```

RESULT 10
US-10-750-185-42261

```



```

Sequence 42261, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMT GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFIELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFE
FILE REFERENCE: MM11100-2
CURRENT APPLICATION NUMBER: US/10750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 42261
LENGTH: 1661
TYPE: DNA
ORGANISM: Bovine
US-10-750-185-42261 19866881270808

```

[illegible]

```

RESULT 11
US-10-860-501-2
; Sequence 2, Application US/10860501
; Publication NO. US20050271638A1
; GENERAL INFORMATION:
;
; APPLICANT: Li, Linheng
; APPLICANT: He, Xi
; TITLE OF INVENTION: BMP PATHWAY METHODS AND COMPOSITIONS
; FILE REFERENCE: 64991
; CURRENT APPLICATION NUMBER: US/10/860,501
; CURRENT FILING DATE: 2004-06-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
;
; LENGTH: 1891
;
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-860-501-2

```

Query Match	14.8%	Score 28.2;	DB 6;	Length 1891;
Best Local Similarity	36.4%;	Pred. No. 5.3;		
Matches	44;	Conservative 19;	Mismatches 58;	Indels 0; Gaps 0;
Oy	18	CUUCGCGAGCGAAAAUUNUGCACAGUAUUAAAUNCUSCAGUAGUGGUUAUNUGUGAAUAC	77	
Db	1211	CCTGAAGAGCAAAACATCCTTATTTAAGAAAAATGAAAGTTCCTGTATTGCCTGACCCTGGG	1270	
Oy	78	CGUACCUAUNUGAGUUUAUGCUCACAGAUCGUGUGAUNAGGCCCUACAUAUUCUACGA	137	
Db	1271	CCTAGCTOTTAATTCAACACAGTGATCAAAATGAATTGACATACCCCTTGAAATACCAAGGT	1330	
Oy	138	G	138	

Db 1331 G 1331

```

RESULT 12 505501-1
US-10-860-501-1
; Sequence 1, Application US/10860501
; Publication No. US20050271638A1
; GENERAL INFORMATION:
; APPLICANT: LI, Linheng
; APPLICANT: He, Xi
; TITLE OF INVENTION: BMP PATHWAY METHODS AND COMPOSITIONS
; FILE REFERENCE: 64991
; CURRENT APPLICATION NUMBER: US/10/860,501
; CURRENT FILING DATE: 2004-06-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2056
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-860-501-1

```

Query Match	14.8%	Score 28.2;	DB 6;	Length 2056;
Best Local Similarity	36.4%	Pred. No. 5.5;		
Matches 44;	Conservative 19;	Mismatches 58;	Indels 0;	Gaps 0;

Qy	18	CTUCGGAGGCGAAAUUUGCAGAUAAAAUUCGCAAGUAGUCUAUUUGGAAUAC	77
Db	1376	CCTGAMAGGCAAAAAACATCCTTATTAGAAAAATGGAAGTGCTGTATGTGACCTGGG	1435
Qy	78	CGUACCUAUUUAGSUUUACGUCCAAGUUGGUGAUAAGCAGCCUUNCAUAUUCUAGGA	137
Db	1436	CCTTAGCTGTAAATTCACAAGTGATACAAATGAAGATTGACATACCTTGAATACCAAGGT	1495
Qy			
		138 G 138	
Db		1496 G 1496	

```

RESULT 13
US-10-750-185-48297
Sequence 48297, Application US/10750185
Publication No. US2005026063A1
GENERAL INFORMATION:
APPLICANT: MMT GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM11100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 48297
LENGTH: 1946
TYPE: DNA
ORGANISM: Bovine
US-10-750-185-48297 19866890368423

```

```

Query Match Similarity 14.6%; Score 27.8; DB 6; Length 1946;
Best Local Similarity 34.5%; Pred. No 7.4;
Matches 41; Conservative 21; Mismatches 57; Indels 0; Gaps 0
Oy 66 UGUGAUAUCCGUACCUAUAUUAGUUUACGCCUCAAAGUUGGUGUAGACGCCUUAU 125
Db 144 TGGTAGTAGAAACCTGGCTAATTAATTGTTATCTTCAACCTGTAGAAACAGCCCTC 203

```


QY 126 CAAUAUCUAGAGAACAUCUGCUAUGUUUUAAGAAGUUAGUACUCAAACAGAACAU 184
 : | : : : : : : : : : : : : : : : : : :
Db 204 TTCGCCATCCCAATCTGTACTTTGGGGAATAATTATTAGTTAATTACAGATTATT 262

RESULT 14

```

US-11-092-140-101
; Sequence 101, Application US/11092140
; Publication No. US20050262590A1
;
GENERAL INFORMATION:
;
APPLICANT: Subramaniam, S.; Slater, S.; Karberg, K.; Chen, R.; Valentin, H.; Wong, Y
;
FILE OF INVENTION: Nucleic Acid Sequences to Proteins Involved in Tocopherol Synthes
;
TITLE REFERENCE: 16515.054
;
CURRENT APPLICATION NUMBER: US/11/092,140
;
CURRENT FILING DATE: 2005-03-29
;
PRIOR APPLICATION NUMBER: US/09/688,069
;
PRIOR FILING DATE: 2000-10-14
;
NUMBER OF SEQ ID NOS: 114
;
SEQ ID NO 101
;
LENGTH: 705
;
TYPE: DNA
;
;
ORGANISM: Allium porrum
;
US-11-092-140-101

```

Query Match	14.5%	Score 27.6;	DB 7;	Length 705;
Best Local Similarity	36.7%;	Pred. No. 5.7;		
Matches 36;	Conservative 18;	Mismatches 44;	Indels 0;	Gaps 0;

Oy 29 AAAAUAUUGCAGCUAUAAAAUCUGCAAGUAGUCUAUUGUGGAAUACCCGUACCUAUUU 88
| | : | : | : | : | : | : | : | :
Db 142 ACAACATGAATTATAATTAAAGAGAAGAAATGCATCTGCAATGTGGTAAAAGAACTCTCCGTTTTT 201

Dy 89 AGGUTUACCGUCCAAGAUCCSGUGAUAAGCACCCTUAC 126
| : | : | : | : | : | :
Db 202 AAGATGCACATTAAAGCGTTGTTAATTCCAGCCATTTC 239

RESULT 15

```

?      Sequence 61632, Application US/10750185
?      Publication No. US20050260603A1
?      GENERAL INFORMATION:
?      APPLICANT: MMT GENOMICS, INC.
?      APPLICANT: DENISE, Sue K.
?      APPLICANT: KERR, Richard
?      APPLICANT: ROSENFIELD, David
?      APPLICANT: HOLM, Tom
?      APPLICANT: BATES, Stephen
?      APPLICANT: FANTIN, Denis
?      TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
?      FILE REFERENCE: MW1100-2
?      CURRENT APPLICATION NUMBER: US/10/750,185
?      CURRENT FILING DATE: 2003-12-31
?      PRIOR APPLICATION NUMBER: US 60/437,482
?      PRIOR FILING DATE: 2002-12-31
?      NUMBER OF SEQ ID NOS: 64922
?      SOFTWARE: Patencin version 3.1
?      SEQ ID NO 61632
?      LENGTH: 1223
?      TYPE: DNA
?      ORGANISM: Bovine
?      US-10-750-185-61632

```

Query Match	14.5%	Score 27.6;	DB 6;	Length 1223;
Best Local Similarity	33.6%;	Pred. No. 7.1;		
Matches 49;	Conservative 23;	Mismatches 74;	Indels 0;	Gaps 0;

[illegible]

Db 645 AGTATTCGCGAAACTTCTCAGAAAAACAATATTAAGTTGAAGTCAAGCATTTT 586
 Qy 150 GUTUAGAAVUAGUAGUCUCUAAA 175
 Db 585 GATCAAAAAGATATATGATTTTTTAAA 560

Search completed: December 22, 2005, 16:41:12
Job time : 176.374 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 04:46:32 ; Search time 1380.53 Seconds
(without alignments)
7205.664 Million cell updates/sec

Title: US-10-088-750C-7

Perfect score: 175
Sequence: 1 aguuugugugauucugcgc.....acuaucuuauauuagaccu 175

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Genembl: *
1: gb_ba: *
2: gb_in: *
3: gb_env: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pr: *
9: gb_ro: *
10: gb_sbs: *
11: gb_gy: *
12: gb_un: *
13: gb_vl: *
14: gb_mtg: *
15: gb_pl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	175	100.0	175	6	BD173517 Novel ter
2	175	100.0	175	6	BD177021 Novel tra
3	175	100.0	10011	13	AF022937 Rhopalosi
4	36.8	21.0	91923	8	AC108007 Homo sapi
5	36.8	21.0	160494	8	AC055860 Homo sapi
6	36.8	21.0	172939	8	AC090099 Homo sapi
7	35.2	20.1	9812	13	AF535531 Aphid let
8	35.2	20.1	9829	6	AX763949 Sequence
9	35.2	20.1	240387	14	AC107353 Rattus no
10	35.2	20.1	272055	14	AC118859 Rattus no
11	34.8	19.9	110000	15	AP008209_158
12	34.8	19.9	152423	15	AC079853 Oryza sat
13	34.4	19.7	57053	14	CR383679 Danio rer
14	34.4	19.7	151303	5	CR388405 Zebrafish
15	34.2	19.5	25935	5	BA000027_4
16	33.6	19.2	156371	14	CR847526 Danio rer
17	33.6	19.2	223522	14	CR759920 Danio rer
18	33	18.9	378	10	GA6118 Z5563_1 Zeb

c	19	33	18.9	38396	14	AY714850	AY714850 Unculture
c	20	33	18.9	110000	15	AP008218_029	Continuation (30 o
c	21	33	18.9	145661	15	CNS08CNU	AL844498 Oryza sat
c	22	33	18.9	194001	14	AC151251	AC151251 Bos tauru
c	23	33	18.9	222428	14	AC150960	AC150960 Bos tauru
c	24	32.6	18.6	225911	14	CR925880	CR925880 Danio rer
c	25	32.6	18.6	279428	14	CR854891	CR854891 Danio rer
c	26	32.2	18.4	177531	5	BX005396	BX005396 Zebrafish
c	27	32	18.3	110000	14	BX927305_2	Continuation (3 of
c	28	32	18.3	188167	5	BX890568	BX890568 Zebrafish
c	29	32	18.3	239986	14	AC107150	AC107150 Rattus no
c	30	32	18.3	260006	14	AC128958	AC128958 Rattus no
c	31	32	18.3	312430	1	RFX02	AJ355271 Rickettsi
c	32	31.8	18.2	1140	5	AY281055	AY281055 Poxvirus
c	33	31.8	18.2	202971	14	CR759948	CR759948 Danio rer
c	34	31.8	18.2	205742	14	CR376742	CR376742 Danio rer
c	35	31.8	18.2	268294	14	AC155761	AC155761 Bos tauru
c	36	31.6	18.1	5853	6	AX283321	AX283321 Sequence
c	37	31.6	18.1	161393	9	AL645966	AL645966 Mouse DNA
c	38	31.6	18.1	201500	9	AL645807	AL645807 Mouse DNA
c	39	31.4	17.9	163093	5	BX682544	BX682544 Zebrafish
c	40	31.4	17.9	201429	14	AC109375	AC109375 Rattus no
c	41	31.4	17.9	256796	14	AC099183	AC099183 Rattus no
c	42	31.2	17.8	189	6	BD173513	BD173513 Novel ter
c	43	31.2	17.8	189	6	BD177017	BD177017 Novel tra
c	44	31.2	17.8	4158	1	AB002529	AB002529 Pseudomon
c	45	31.2	17.8	9264	13	AF014388	AF014388 Drosophila

ALIGNMENTS

RESULT 1	BD173517	175 bp	RNA	linear	PAT 18-FEB-2003
LOCUS	Novel tertiary structure having ability to accelerate translation activity.				
DEFINITION	Novel tertiary structure having ability to accelerate translation activity.				
ACCESSION	BD173517.1 GI:28414848				
VERSION	BD173517.1 GI:28414848				
KEYWORDS	WO 02061080-A/7.				
SOURCE	Rhopalosiphum padi virus				
ORGANISM	Rhopalosiphum padi virus				
REFERENCE	1 (bases 1 to 175)				
AUTHORS	Nakashima,N. and Kanamori,Y.				
TITLE	Novel tertiary structure having ability to accelerate translation activity.				
JOURNAL	Patent: WO 02061080-A 7 08-AUG-2002;				
COMMENT	JAPAN AS REPRESENTED BY DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND ENTOMOLOGICAL SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND FISHERIES, NOBUHIKO NAKASHIMA, YASUSHI KANAMORI				
OS	Rhopalosiphum padi virus				
PN	WO 02061080-A/7				
PD	08-AUG-2002				
PR	31-JAN-2001 WO 2001JP000641				
PI	25-JAN-2001 JP 01P 016746				
PC	NOBUHIKO NAKASHIMA, YASUSHI KANAMORI				
CC	C12N15/11, C12N15/86, C12P21/02				
FT	Novel tertiary structure having ability to accelerate translation activity				
FT	Key				
FT	Location/Qualifiers				
FT	source				
FEATURES	Location/Qualifiers				
source	1..175				
	/organism="Rhopalosiphum padi virus"				
	/mol_type="genomic RNA"				
	/db_xref="taxon:66834"				
ORIGIN					
Query Match	100.0%; Score 175; DB 6; Length 175;				
Best Local Similarity	71.4%; Pred. No. 3.1e-50;				

Matches 125; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGUGUGUGUAGUCUGCGCGCAUUAAGUCUGGAGAAAGCGUUGCCUUAUUGCUACAACA 60
Db 1 AGTGTGTGTGTATCTTGCGCGCAATAATGCTGACGTAATGCTGATGCTGACCA 60

QY 61 CUGGUGUACUUAUUGUCUUAUCUAUACAGACCGCGUGUGACGCCACCAAAAGUCUAG 120
Db 61 CTGAGTTAGCTTATTAAGCTTACTTAATCAAGACGCGCTGCTGACGCCACCAAAAGCTTAG 120

QY 121 AUAAGUCAACAGAGAGCAUACGUGUGCGUGUGUACUUAUUGCUUAUUAUAGCCU 175
Db 121 ATACGTCACAGAGAGCATACGCTAGTGCCTGACTATCTTATATATGACCT 175

RESULT 2
BD177021 175 bp RNA linear PAT 16-APR-2003
LOCUS Novel translational activity-promoting higher-order structure.
DEFINITION BD177021.1 GI:30014281
ACCESSION UP 2002306168-A/7.
VERSION Rhopalosiphum padi virus
KEYWORDS Rhopalosiphum padi virus
SOURCE Viruses; ssRNA positive-strand viruses, no DNA stage;
ORGANISM Dicistroviridae; Crispavirus.
REFERENCE 1 (bases 1 to 175)
AUTHORS Nakashima, N. and Kanamori, Y.
TITLE Novel translational activity-promoting higher-order structure
JOURNAL Patent: JP 2002306168-A 7 22-OCT-2002;
DIRECTOR GENERAL OF NATIONAL INSTITUTE OF SERICULTURAL AND
ENTOMOLOGICAL, HIROSHIMA PREFECTURAL INDUSTRIAL TECHNOLOGY PROMOTION
ORGANIZATION, SCIENCE MINISTRY OF AGRICULTURE FORESTRY AND
FISHERIES

COMMENT OS Rhopalosiphum padi virus
PN JP 2002306168-A/7
PD 22-OCT-2002
PF 25-JAN-2001 JP 2001016746
PI NOBUHIKO NAKASHIMA, YASUSHI KANAMORI
PC C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02// PC
(C12N15/09, C12R1/92), C12N15/00, C12N5/00, C12N15/00, C12R1/92) CC
Novel translational activity-promoting higher-order structure PH
Key Location/Qualifiers
FT source 1..175
FT location/Qualifiers
ORIGIN 1..175
location/Qualifiers
source 1..175
/organism="Rhopalosiphum padi virus"
/mol_type="genomic RNA"
/db_xref="taxon:66834"

Query Match 100.0%; Score 175; DB 6; Length 175;
Best Local Similarity 71.4%; Pred. No. 3.1e-50;
Matches 125; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGUGUGUGUAGUCUGCGCGCAUUAAGUCUGGAGAAAGCGUUGCCUUAUUGCUACAACA 60
Db 1 AGTGTGTGTGTATCTTGCGCGCAATAATGCTGACGTAATGCTGATGCTGACCA 60

QY 61 CUGGUGUACUUAUUGUCUUAUCUAUACAGACCGCGUGUGACGCCACCAAAAGUCUAG 120
Db 61 CTGAGTTAGCTTATTAAGCTTACTTAATCAAGACGCGCTGCTGACGCCACCAAAAGCTTAG 120

QY 121 AUAAGUCAACAGAGAGCAUACGUGUGCGUGUGUACUUAUUGCUUAUUAUAGCCU 175
Db 121 ATACGTCACAGAGAGCATACGCTAGTGCCTGACTATCTTATATATGACCT 175

RESULT 3
AF022937 10011 bp RNA linear VRL 16-DEC-1998
LOCUS Rhopalosiphum padi virus complete genome.
DEFINITION

ACCESSION AF022937
VERSION AF022937.1 GI:2911298
KEYWORDS Rhopalosiphum padi virus
SOURCE Rhopalosiphum padi virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage;
REFERENCE 1 (bases 1 to 10011)
AUTHORS Moon, J.S., Domier, L.L., McCoppin, N.K., D'Arcy, C.J. and Jin, H.
TITLE Nucleotide sequence analysis shows that Rhopalosiphum padi virus is a member of a novel group of insect-infecting RNA viruses
JOURNAL Virology 243 (1), 54-65 (1998)
PUBMED 9527915
REFERENCE 2 (bases 1 to 10011)
AUTHORS Domier, L.L., Moon, J.S., McCoppin, N.K. and Jin, H.
TITLE Direct Submission
JOURNAL Submitted (05-SEP-1997) Crop Sciences, University of Illinois, USDA-ARS-MMA-CPRU, 1102 S. Goodwin Ave., Urbana, IL 61801, USA

FEATURES
source location/Qualifiers
1..10011
/organism="Rhopalosiphum padi virus"
/mol_type="genomic RNA"
/db_xref="taxon:66834"
/note="RHPV"
580..6576
/codon_start=1
/product="nonstructural polypeptidein"
/protein_id="AAC95509.1"
/translation="MSTMSCTNLPKSVQRIQHCNSLASEBRLVTLPRKDFSTIDSP
KHLIDLPFSEVGNRSEVCDNEPVLGYKSTLIEQMFIRKYLKFNFEERQSLIAK
RMBEIRLEFFRSGRTIYPLNMYPLFQDTNSIRHPANYRWMGSEYLDGDIKRFPL
ILMFCINIEBVPURCYKFSANVMKVALSOLPLMALVSTTEYVLGDMIDPKYVANSK
HFRFRLSYTNPIDLEHSTHELGLYCALCWIRKIPSTTEYVLGDMIDPKYVANSK
ENMDVYRLQTVAQLDGENTVQMDPAFQALDKADTGEAVTSLHSAVAVDGI
RKRFACVLSICNLYRLSGSGMSQDVLNMLVTTLQSDMPANIIPQLRSIPLTSTAQ
SASIDANVITKLALCSFSLMSRIPPSRIDFTIILDRIPKAFLENIWKRDLIV
TNELMTWMEVTVLKRENVIPRSIDLSGSEKNDLBOLTLQKREIQTNLSTQHAG
RMVSEGRILMRICKNDLSKNGTEITARNPAKLLTEANMSGADSKIRTEPVIVM
FSGASNGKGLGSEYPLDMMRYVGPDPDQOONVYAREEETWQXINOEYIVVD
FIOTKPSQKNEBLEBMRIRLGNMFPYQCHMASLIDKNIPAEKLCITSLNLRDIT
ESLNCSPAVSRRIIDFAFNRIIPBYQMEVSNAGDKYRLDPAKARQDFEDVLCFEYV
RFLDPFASRSRNLITDLTYTEVVKLCDDKRRKADNANNEDVLALESYRNKGVAKSEK
KHETDSSYNGEITVFTSTAGVHLDDVNLVILQPSYKRLYNNIKCKYSTRMLAST
DMSAEFLMLRAGDGAIDKCLSYRETRKSLNSINRETVIKNVFENYVLPFKACG
AATGASVYLIIRKKTPTPTTAAPVNGKELLTTRKNECECDNCRNKKKHKHNVDC
VKNYTRCHCYALHMEKAQVNLKYTTAAATQEBEMKREBRCVBLITIDQLSCDCA
NCDACCDDSLTERFENVMKYDMPCVVCVCAKLSQGFPMVBLALIKRGLKEPPLIN
PYLRKLSVKLSIEDIRDEFERTTDHDLNLTQSGEYBEBVTKIIRKVALRYGSHDDT
NMRLRKILPRVKYQTESVDVLSVSGPTKSGSDIDKIIDVNNIEQGMEMDSKSVT
IVNHVVYPTNVMTANKNDGTETNIGIIFIICGVVLMPPHYKLAIEBROGYSNLVS
RKLIQAKIPSSVFDTPVRLQGDAMLVAFPTVNSFKNIVSHFVDIOSYLVIVSGPDI
LAKYRNAMTEBSRVCISALIGVSEBDEVDVMSVPCGMEVYRBDPTTYATPRPAGC
GALCVANITIGQKILGIVHSVGGJGCGKSSATTKOMRESLKQSPSIQVAPSSGA
LTYEMDLAESGSEFVLAHLYLPGVSIQITWQYIAKPIHQQLATPKKPELIDFKIR
GEIILDRILLQKKYGRPYLDONIVNDJBDGLKSVYQSEHEPEVEYKPIIFDQAI
LGIDGDPFISLDRNTAPGEPSTMRGDTKGLMENGMDYDLTGYYAVALADVDY
LENDINGTRPELIVMTDITLKDOKIAIAKANAGTRTFSAAPMRYAIALRYVCAPVNH
LSMRIRNNTCYGVNPPSSSWSAVAKOLKYKSHVIAAGVSNDCGSI.PADLYAAETI
MADYRIDLMDYVAHAKNAVVGQVLLDEFLDILRLRYIECVHHLHMANRQSSLMYI
VRNGIPSGCVPYAPLNSIVNQMALVYCWNIIDDPILRNKVEFEFHTSIYFYDDFWM
NIRADVLEQNOQTTITRAMSDYLDMMVTEBAKGECKVMTLEBVFLEKCAFYNTEFI
OEYTAAPDLIVLIDSNTMYKIGCSAVIVARDTLKACLRALAHPEHIDSOYSIITE
IGLQILNLIPGELFIPDTRYSTLAIKNNMSDGLGALNCDA"

CDS
join(7103..7105,7107..9560)
/note="The site of translation initiation could not be determined from the nucleotide sequence, but may involve a +1 frameshift"
/codon_start=1
/product="structural polypeptidein"
/protein_id="AAC95510.1"
/db_xref="GI:2911300"
/translation="MPANINENTTYKIQQOILSFSSBGBSPSSSTVLAPLKLQDPILID

CARDKTHVNSFLERPIINPRATWNSOPAGEBFLSFNYPENNVYINPMYSKLNPLG
LRADLVYRVQVNAOBFAHAGRLMISWTPRLDYICRNGKYYRSPDLVSCGNTISGPNRVE
IDLSTTEBATPTTFPSFPLLYNNLTGSSGDIOTROLIYSPDLVSCGNTIDITYINW
MTNVRTEPTMPTPSIAQVSEGSTQOQKGVSTQKSEYVSTIMEPLTKIPVGOULIG
AKSCVDALHAVASHTGMSKPLNPADMOJFKQAPSKFMCNSGDMATNLGLTQNLIE
HLQSLFRSDSEMSIDVARTYNNYSRNNIKNNGPGVLTLSHVSPSPTFVETG
LSIHLVYASNPFVLMRGNNIKLKFKVTKFHSGRVRLIYVPGSRKXIAIPATLPTNFRDAN
KSTVVDLRSDPTDVERNNPVYATVPMIYHNSRTPWYATSOIHCASSTIVEVNEIWNIS
TVSTTIEVLVECADEDIERAIPYVPSITPPAPRQGSRRATYDLITSIAVGTDTGD
TPLEVSREPTTENDVPLQPTTTFNASMLMGEKVISFROLIKRFSATIPPTQNRW
EPTQPMINTRLEGTQDESSDIDGISMFASLIAPYNGSRKXIAIPATLPTNFRDAN
KPNLSGAVRIVDTNGTWTYDYGKAEVFMTPNENGIHELSPYSSSYVITLTYNSTD
SDVLDANGFVRVIAFHSPTSAVYKAGDFFSGFLGAPVITIGLSSVRLICVKY
HVRVSPLENTSDMRKLSNE"

ORIGIN

Query Match 100.0%; Score 175; DB 13; Length 10011;
Best Local Similarity 71.4%; Pred. No. 3.6e-50;
Matches 125; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGUGUUGUAGUACUUGCGCGAUAUAGCUGAAGCAAAAGCUGUUCGUAACACACA 60
+:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 6935 AGTGTGTGTGATCTTGCGCATTAATGCTGACGTAAGAAAGCTTGCTATTCACACA 6994
+:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 61 CUUGGUGUAGUAGUACUUAUAGACGCGCGUGGAGCCGACCAAAAGUCUAG 120
+:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 6995 CTTGGTTAGCTATTGATTAAGCTTACTATCAAGACGCGCGTGACCCACAAAAGCTTAG 7054
+:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 121 AUAAGUCACAGAGAGACUAGCUGUAGCGGUGUAGCUUAGUUAUAGACCU 175
+:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 7055 ATAGCTCACAGAGAGCATAGCTAGTGCAGTTACATCTTATATATGACCT 7109
+:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 4
AC108007/c 91923 bp DNA linear PRI 03-SEP-2002
LOCUS Homo sapiens chromosome 11, clone CTD-30381L2, complete sequence.
AC108007
AC108007.5 GI:22655822
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 91923)
Birtten,B., Nusbbaum,C. and Lander,E.
JOURNAL
REFERENCE 2 (bases 1 to 91923)
Birtten,B., Linton,L., Nusbbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barina,N., Baetien,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepey,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Grinde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marguis,N., Matthews,C., McCarthy,M.,
McGowan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mingra,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Plerre,R., Rise,C., Rogov,P., Roman,J.,
Roretz,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strausen,P., Subramanian,A., Talamas,J., Teefaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL
REFERENCES
AUTHORS

Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 91923)
Birtten,B., Nusbbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barina,N., Baetien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepey,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mingra,V.,
Mingra,V., Murphy,T., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Peterson,K., Phunkhang,P., Plerre,R., Rise,C., Rogov,P., Roman,J.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Teefaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCES
AUTHORS
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 91923)
Birtten,B., Nusbbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barina,N., Baetien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepey,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mingra,V.,
Mingra,V., Murphy,T., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Peterson,K., Phunkhang,P., Plerre,R., Rise,C., Rogov,P., Roman,J.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Teefaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT
Submitted (03-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 3, 2002 this sequence version replaced gi:22474954.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: I24596
Center clone name: 3038_L_12

Only the first 91.9 kb of this clone are being submitted.
The remainder overlaps accession number AC055860 (WICGR project
L9785).
----- Location/Qualifiers
1. 91923
/organism="Homo sapiens"
/mol type="genomic DNA"
/db xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="CTD-30381L2"
/clone_id="CITD2 Human BAC"
repeat_region 25..70

repeat_region	/rpt_family="L12" complement(251. ,899) /rpt_family="L12" 900. ,1001
repeat_region	/rpt_family="L1MD2" 1082. ,1655 /rpt_family="L1MD2" 1671. ,1850
repeat_region	/rpt_family="L1MD3" complement(1856. ,1917) /rpt_family="L12" 1918. ,2273
repeat_region	/rpt_family="THE1B" complement(12274. ,2327) /rpt_family="L12" 4505. ,4713
repeat_region	/rpt_family="L12" 5536. ,5795 /rpt_family="AluSc" 5796. ,5837
repeat_region	/rpt_family="TAAAn" 5838. ,5875 /rpt_family="AluSc" complement(7166. ,7526)
repeat_region	/rpt_family="L1MB4" complement(107116. ,10825) /rpt_family="MIR" 11369. ,11466
repeat_region	/rpt_family="L12" complement(11479. ,11644) /rpt_family="MIR3" 12072. ,12409
repeat_region	/rpt_family="MER2" 12692. ,12875 /rpt_family="LTR33" complement(112876. ,13294)
repeat_region	/rpt_family="WSTA" 13395. ,13552 /rpt_family="LTR33" complement(14619. ,14922)
repeat_region	/rpt_family="L1PB1" complement(115002. ,17751) /rpt_family="L1PB1" 17752. ,19390
repeat_region	/rpt_family="L1PA6" 19591. ,20184 /rpt_family="L1PA4" 20185. ,24476
repeat_region	/rpt_family="L1PA6" 24611. ,24637 /rpt_family="Tn" complement(24639. ,26213)
repeat_region	/rpt_family="SVA" 26219. ,26361 /rpt_family="L1MB5" 26372. ,26560
repeat_region	/rpt_family="FLAM_C" complement(26636. ,26719) /rpt_family="MIR" complement(26663. ,27139)
repeat_region	/rpt_family="MER41C" complement(12797. ,28030) /rpt_family="MER41C" complement(127140. ,27449)
repeat_region	/rpt_family="AluBg" complement(127450. ,27494) /rpt_family="MER41C" complement(127495. ,27796)
repeat_region	/rpt_family="AluBg" complement(12797. ,28030) /rpt_family="MER41C" complement(128194. ,28287)
repeat_region	/rpt_family="MIR" complement(129002. ,29304) /rpt_family="AluBo" complement(129002. ,29304)

REFERENCE	TITLE	JOURNAL	AUTHORS
1	repeat_region	complement(29337..29755)	/rpt_family="WISTA"
2	repeat_region	30412..30452	/rpt_family="TG)n"
3	repeat_region	complement(31075..31197)	/rpt_family="MIR"
4	repeat_region	31198..31246	/rpt_family="CA)n"
5	repeat_region	complement(31247..31341)	/rpt_family="MIR"
6	repeat_region	31541..31821	/rpt_family="A1USX"
7	repeat_region	31822..31850	/rpt_family="GA)n"
8	repeat_region	32030..32230	/rpt_family="MERSA"
9	Query Match	21.0%; Score 36.8; DB 8; Length 91923;	
10	Best Local Similarity	46.9%; Pred. No. 0.21;	
11	Matches	30; Conservative 17; Mismatches 17; Indels 0; Gaps 0;	
12	Db	65230 CCAT 65227	
13	QY	84 UAAU 87	
14	Db	65230 CCAT 65227	
15	RESULT 5		
16	AC055860/c	160494 bp	DNA linear PRI 28-DEC-2001
17	LOCUS	AC055860	Homo sapiens chromosome 11, clone RP11-195023, complete sequence.
18	DEFINITION	AC055860	Homo sapiens chromosome 11, clone RP11-195023, complete sequence.
19	ACCESSION	AC055860.13	GI:17985932
20	VERSION	AC055860.13	GI:17985932
21	KEYWORDS	HTG.	
22	SOURCE	Homo sapiens (human)	
23	ORGANISM	Homo sapiens	
24	REFERENCE	Anderson, S., Baldwin, J., Barna, N., Baetien, V., Bede, F., Boguslavsky, L., Boukngalter, B., Brown, A., Burke, G., Campoliano, A., Casale, A., Choell, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeAtellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Gajanan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karatas, A., Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McSheeters, R., Meldrum, J., Menue, L., Mihova, T., Miranda, C., Mlenka, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tasey, S., Theodore, J., Tittrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.	
25	REFERENCE	Anderson, S., Baldwin, J., Barna, N., Baetien, V., Bede, F., Boguslavsky, L., Boukngalter, B., Brown, A., Burke, G., Campoliano, A., Casale, A., Choell, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeAtellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Gajanan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karatas, A., Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McSheeters, R., Meldrum, J., Menue, L., Mihova, T., Miranda, C., Mlenka, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tasey, S., Theodore, J., Tittrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.	
26	REFERENCE	Anderson, S., Baldwin, J., Barna, N., Baetien, V., Bede, F., Boguslavsky, L., Boukngalter, B., Brown, A., Burke, G., Campoliano, A., Casale, A., Choell, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeAtellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Gajanan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karatas, A., Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McSheeters, R., Meldrum, J., Menue, L., Mihova, T., Miranda, C., Mlenka, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tasey, S., Theodore, J., Tittrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.	
27	REFERENCE	Anderson, S., Baldwin, J., Barna, N., Baetien, V., Bede, F., Boguslavsky, L., Boukngalter, B., Brown, A., Burke, G., Campoliano, A., Casale, A., Choell, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeAtellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Gajanan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karatas, A., Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McSheeters, R., Meldrum, J., Menue, L., Mihova, T., Miranda, C., Mlenka, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tasey, S., Theodore, J., Tittrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.	
28	REFERENCE	Anderson, S., Baldwin, J., Barna, N., Baetien, V., Bede, F., Boguslavsky, L., Boukngalter, B., Brown, A., Burke, G., Campoliano, A., Casale, A., Choell, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeAtellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Gajanan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karatas, A., Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McSheeters, R., Meldrum, J., Menue, L., Mihova, T., Miranda, C., Mlenka, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tasey, S., Theodore, J., Tittrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.	
29	REFERENCE	Anderson, S., Baldwin, J., Barna, N., Baetien, V., Bede, F., Boguslavsky, L., Boukngalter, B., Brown, A., Burke, G., Campoliano, A., Casale, A., Choell, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeAtellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., G	

Anderson, S., Barna, N., Baetien, V., Boguslavskiy, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Chepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArliano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gaidyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, N., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Labroque, K., Lamezarez, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margis, N., Mathews, C., McCarthy, M., McEwan, P., McKernan, K., McSheehy, R., Meldrum, J., Menous, L., Mihova, T., Mienna, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retter, R., Riebeck, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (28-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 28, 2001 this sequence version replaced gi:17223377.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 19785
Center clone name: 195_O_23

FEATURES
source

Location/Qualifiers
1.160494
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-195023"
/clone_lib="RP11 Human Male BAC"
complement(1..213)
repeat_region
/rpt_family="MER41C"
complement(214..523)
repeat_region
/rpt_family="AluSg"
complement(524..568)
repeat_region
/rpt_family="MER41C"
complement(569..869)
repeat_region
/rpt_family="AluSg"
complement(870..1103)
repeat_region
/rpt_family="MER41C"
complement(1127..1360)
repeat_region
/rpt_family="MIR"
complement(1205..2377)
repeat_region
/rpt_family="AluJo"
complement(2410..2828)
repeat_region
/rpt_family="MSTA"
3485..3521
repeat_region
/rpt_family="(TG)n"
complement(4144..4266)
repeat_region
/rpt_family="MIR"
4267..4315
repeat_region
/rpt_family="(CA)n"
complement(4316..4410)
repeat_region
/rpt_family="MIR"
4610..4890
repeat_region
/rpt_family="AluSx"
4891..4919

repeat_region
/rpt_family="(GA)n"
5099..5299
repeat_region
/rpt_family="MER5A"
complement(5719..5816)
repeat_region
/rpt_family="MIR"
7120..7232
repeat_region
/rpt_family="L2"
7459..7653
repeat_region
/rpt_family="L1M4"
7691..7810
repeat_region
/rpt_family="AluJo"
7811..7841
repeat_region
/rpt_family="(CAA)n"
7842..8083
repeat_region
/rpt_family="AluJo"
complement(8766..8957)
repeat_region
/rpt_family="MIR"
complement(9605..9915)
repeat_region
/rpt_family="AluJb"
complement(10266..10552)
repeat_region
/rpt_family="AluSx"
10576..10610
repeat_region
/rpt_family="AT-rich"
10740..10997
repeat_region
/rpt_family="AluY"
11001..11029
repeat_region
/rpt_family="AT-rich"
11036..11159
repeat_region
/rpt_family="FLAM-C"
11162..11516
repeat_region
/rpt_family="MER65C"
11524..11629
repeat_region
/rpt_family="MER20B"
11630..11945
repeat_region
/rpt_family="AluY"
11946..12060
repeat_region
/rpt_family="MER20B"
13219..13317
repeat_region
/rpt_family="AluJ/FLAM"
14235..14533
repeat_region
/rpt_family="AluJo"
14557..14722
repeat_region
/rpt_family="MIR"
14744..15124
repeat_region
/rpt_family="L2"
complement(115461..16275)
repeat_region
/rpt_family="MER6"
complement(16920..17030)
repeat_region
/rpt_family="MIR"
17047..17207
repeat_region
/rpt_family="L2"
complement(117232..17293)
repeat_region
/rpt_family="L1ME1"
complement(17294..17587)
repeat_region
/rpt_family="AluSx"
17310..17314
repeat_region
/note="<30 qual SNGL region"
17378..17382
repeat_region
/note="<30 qual SNGL region"
17412..17416
repeat_region
/note="<30 qual SNGL region"
complement(17588..18054)
repeat_region
/rpt_family="L1ME1"
complement(18055..18429)
repeat_region
/rpt_family="THEIC"
complement(18430..18991)
repeat_region
/rpt_family="L1ME1"
complement(19131..19256)
repeat_region
/rpt_family="AluJb"
complement(19309..19563)
repeat_region
/rpt_family="L1ME1"
19872..20070
repeat_region
/rpt_family="L2"

REFERENCE	TITLE	JOURNAL	AUTHORS
AC090009/c	LOCUS	AC090009/c	
AC090009	DEFINITION	Homo sapiens chromosome 11, clone RP11-113D6, complete sequence.	
AC090009	VERSION	AC090009.14	GI:23396332
AC090009	KEYWORDS	HTG.	
AC090009	SOURCE	Homo sapiens (human)	
AC090009	ORGANISM	Homo sapiens	
AC090009	REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Carcharia; Homnidae; Homo.	
AC090009	TITLE	1 (bases 1 to 172939)	
AC090009	JOURNAL	2 (bases 1 to 172939)	
AC090009	REFERENCE	Unpublished	
AC090009	AUTHORS	Blumen, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, S., Bama, N., Besten, V., Boguslavsky, L., Bouhagiel, B., Brown, A., Camarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Haggos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karata, A., Lacroque, K., Lamazares, R., Landers, T., Lehocck, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollars, V., Raymond, C., Rector, R., Riebeck, M., Riley, R., Riese, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Sougnez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	
AC090009	REFERENCE	Submitted (14-FEB-2001) Whitehead Institute/MIT Center for Genome Research, 330 Charles Street, Cambridge, MA 02141, USA	
AC090009	REFERENCE	3 (bases 1 to 172939)	
AC090009	REFERENCE	Blumen, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Bama, N., Besten, V., Bloom, T., Boguslavsky, L., Bouhagiel, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,	

TITLE	Direct Submission
JOURNAL	Submitted (05-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	4 (bases 1 to 172939)
AUTHORS	Barren,B., Nusbach,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barina,N., Basclett,V., Bloom,T., Boguslavsky,L., Boulghalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collimore,A., Cook,A., Cooke,P., Dekrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Farko,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karacas,A., Kellis,C., MacDonald,P., Major,J., Matthews,C., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarty,M., Meldrim,J., Menus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunhthang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tsafaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Vei1,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
FEATURES	Location/Qualifiers 1..172939 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="11" /map="11" /clone="RP11-113D6" /clone_1lb="RP11-11 Human Male BAC" /rpt_family="A1US" complement(1..77) /rpt_family="A1US" complement(154..680) /rpt_family="Char11el" complement(678..810) /rpt_family="Char11el" complement(879..1057) /rpt_family="L1M4" complement(1395..1513) /rpt_family="FLM_A" 1870..1876 /note="<30 qual SNG1 region"
COMMENT	Genome Center Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu ----- Project Information Center project name: L13764 Center clone name: L13_D_6 ----- Direct Submission Submitted (01-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Oct 1, 2002 this sequence version replaced gt:22726159. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996:1197) http://ftp.genome.washington.edu/RM/RepeatMasker.html
JOURNAL	
TITLE	


```
unure      1902..1910
/notes="<30 qual SNGL region"
repeat_region 2672..2788
/rpt_family="SS"
repeat_region complement(4009..4170)
/rpt_family="MIR"
repeat_region complement(4210..4355)
/rpt_family="MERSB"
repeat_region 4356..4647
/rpt_family="AluXs"
repeat_region complement(6554..6658)
/rpt_family="LTR41"
repeat_region 6916..6985
/rpt_family="MLT1C"
repeat_region complement(7175..7359)
/rpt_family="LTR41"
repeat_region complement(8308..8611)
/rpt_family="AluCb"
repeat_region 8622..9083
/rpt_family="MLT2B1"
repeat_region complement(9285..10668)
/rpt_family="L1PB3"
repeat_region 10699..10809
/rpt_family="MERS34"
repeat_region 10782..11264
/rpt_family="LTR48B"
repeat_region complement(11344..11683)
/rpt_family="L2"
repeat_region 11839..12026
/rpt_family="L1MC4a"
repeat_region 12188..12660
/rpt_family="L1MA2"
repeat_region 12879..13222
/rpt_family="L1MC4a"
repeat_region 13323..13522
/rpt_family="AluXs"
repeat_region 13523..13817
/rpt_family="L1MC4a"
repeat_region 14819..15108
/rpt_family="AluO"
repeat_region 15114..15175
/rpt_family="(GAA)n"
repeat_region 16150..16336
/rpt_family="MERSA"
unure      complement(16587..16598)
/notes="<30 qual SNGL region"
unure      complement(16607..16612)
/notes="<30 qual SNGL region"
repeat_region complement(16746..17360)
/rpt_family="L1ME3"
repeat_region 17361..17664
/rpt_family="AluO"
repeat_region complement(17665..17774)
/rpt_family="L1ME3"
repeat_region complement(17827..17929)
/rpt_family="MIR"
repeat_region 18993..19298
/rpt_family="AluO"
repeat_region 19759..19821
/rpt_family="L2"
unure      19933..19941
/notes="<30 qual SNGL region"
repeat_region 20320..20609
/rpt_family="AluO"
repeat_region complement(20684..20994)
/rpt_family="AluXs"
unure      20713..20721
/notes="<30 qual SNGL region"
unure      20879..20883
/notes="<30 qual SNGL region"
repeat_region complement(21118..21398)
/rpt_family="AluXs"
unure      21141..21147
```

```
unure      /note="<30 qual SNGL region"
21144
/note="probably A; possibly G"
unure      21146
/note="probably A; possibly G"
unure      21153..21158
/note="<30 qual SNGL region"
unure      21154
/note="probably T; possibly C"
unure      21160..21167

Query Match      21.0%; Score 36.8; DB 8; Length 172939;
Best Local Similarity 46.9%; Pred. No. 0.22;
Matches 30; Conservative 17; Mismatches 17; Indels 0; Gaps 0;

Cy      24 AAAGCAGACGUGAAACGUGGCUAUGGUCACACACUGGUAACUAGUACUUAAC 83
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      169834 AAGGCTTACTTGAAACATTACTGATCCACAAACCTGGCCACCTATATCATTTAC 169775

Cy      84 UAAU 87
|||
Db      169774 CCAT 169771

RESULT 7
AF536531      9812 bp      RNA      Linear      VRL 16-DEC-2002
LOCUS Aphid lethal paralysis virus, complete genome.
DEFINITION AF536531
ACCESSION AF536531
VERSION AF536531.1 GI:24431318
KEYWORDS Aphid lethal paralysis virus
SOURCE Aphid lethal paralysis virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage;
          Dicistroviridae; Crispavirus.
REFERENCE 1 (bases 1 to 9812)
          Van Munster M., Dulleman A.M., Verbeek M., Van Den Heuvel J.F.,
          Clerivet A. and Van Der Wilk F.
          Sequence analysis and genomic organization of Aphid lethal
          paralysis virus: a new member of the family Dicistroviridae
          J. Gen. Virol. 83 (Pt 12), 3131-3138 (2002)
JOURNAL 12466490
PUBMED 2 (bases 1 to 9812)
AUTHORS van Munster M., van den Heuvel J.F., J.M. and van der Wilk F.
TITLES Direct Submission
JOURNAL Submitted (08-AUG-2002) BioInteractions and Plant Health, Plant
          Research International, P.O. Box 16, Wageningen 6700 AA, The
          Netherlands

FEATURES
source      Location/Qualifiers
1..9812
/organism="Aphid lethal paralysis virus"
/mol_type="genomic RNA"
/db_xref="taxon:209529"
540..6647
/note="encodes putative helicase, protease, and
RNA-dependent RNA polymerase"
/codon_start=1
/product="nonstructural polypeptide"
/protein_id="AAN61470.1"
/translaction="MSGSTFTLLPQLKQSLPBLTQKIRDAQAEKPRIVAOLTRS
TLENMNTYIMTEPLDAIPENTLRKQVLAITEENHNGVSRSAIQAQDNIAVLT
LYLDYMOHTKTAQTCDFPMKAIAMAKKEHFIIDGSPYSIAIPLINASLATI
LNRPDMKPTNLRPRYCLIRNPLKRIISVTLTGILLRKYTIIDPKSAVANNIT
LNAEPRTYIOLIKMTNPNIRNAFDIMTFAYTANFTTDADITSGSSETIOYLN
TLVINKSAIAQGLANFNHNIYVNNIELAINSLTWFAULVNAISDKDETVAYOR
KIVALITINLYRFSSTRKMDLDLVNVGAGLSAGIASLVAOIVTSRAYVAEI
AQIDTSHLLILKALSLTMCCLPIKQPGKTIIDFVTRADRPKAIISGDSMSKID
VVVKQIVMEIKKHNRNNTMDQVDMASEVEKYLELANRNEIKDDPATVLA
SKLYRQVRLIKETQKLSRPNANLIRSLIPATKQADAEIKSGALKQKALREPLIV
WFGKSGMGKTSYSPFMDMKRVFGDIPDFOKNITGRVPEITRYNGTDOETIYD
DAFOIKONVAKPNELEFELIRLGAAPVLMAMSVEKKNTFAPKCVLLTISNDRIK
TESINSPEAVQRRIDPANYNDIAEYREYVDONGNEKYKLNAVKAARAKJGTSTV
```


SNNLDVYRPTKSAFDGRTIETDITYAOTBECSSKMASSRNOHMDPTNYIDAYRNP
 YEPEDIVHPIDAEQIGNAAKVAAGNAGELVTHCIRKYVHGDDTTLIALIRKTK
 RTMENVILDTYGLIPRARDIWEPQSWDHYLYARQIKOYLKQASGIMETLAGFGKW
 QFVKAVALVGVSVFTMVCVKSFALGVYVGGKLVPKYISNDQVAGVSAANAFCPK
 GCNNCKLCKNAMQNLVKNNSSCLYKRMCKGMDINKKASISYSGORSHNHSLS
 VSDMAEIIINOVADCDNCCDCONDELLDLSTVARIHGTNCVCLTRFYOGFPTADL
 YDMTKOJNTLPSITIKNELEKRLIDSHARYVDNAVSANSTRIVGHGATYDARPRR
 MTKIQAOPVTDNRVTIKKNTITLAAQLLPAFAIEPTETKMDIQTPACGDCNS
 PIDQRSLEPDOKSIVIARUVYRNLRFVTVNENETERRRGGQIFMGGGLMI
 PKHPLYMEHVETFGSECHVLENCNALSPLPCSVILIEORHNVDPDRDAIVQ
 LPTNIGFQAQAVKHIIDEQDLARVDSPAIARQNAEBSRQGTNYRREI FMLSTA
 TPEDHIVSTBEANVVVONSGSYTHAVTFEGDGSLLISMAITOKIKMGIMAGI
 THMKGISVALTROIIDRLMKHPOBISOYGHAEIVDLPDPOILKENGTEFLVGEIR
 RIMGSVTRALKESPAVYKGLISSPNKPGTLRPTDENGTTIDPMTIQRKVVARPYIS
 QORVDAVEANSVFYQREYINSFATKLPTRREETIVGINDPPTMALNRTAGPYR
 TYEKESGSGKTKMGDMDYDLTNDACQALMADVELACMLDVRPPIIWIIDLKKA
 KIPILAKAVGKTRFLTACPLHYTLFRQYVFLFLAHARNVONSIAVIGMSPEND
 LLAKRLKNGKHVLIAGYSNFGDGLPVQYVEVAVKIMDPMRMDQIVADBRVING
 RELIDYFEOPFLMKIGVGCINHIHANSDLTGAALYYVRNGI PSCGPATALLNSIV
 NHCCADSWLDMOCTPYATMNSPFEHTSIFYGDDPTMIRPEVIDYNOETITPVIL
 KKNLMTWIDEAKTGEVCAKATLAEVSFLKRFRESVGLMVAIPIDIVLDPNVY
 RIGQOLPLRICVDTLISGGLTELAMHDKVTDAMKMTKMINIGLIDLTREGLEFENDSR
 TTLKLRNEELGSDPEINY"
 <6820..9243
 /note="encodes CP1, CP2 and CP3; start codon not
 determined"
 /codon_start=1
 /product="capsid protein precursor"
 /protein_id="AA061471.1"
 /db_xref="GI:24431320"
 /translation="PANFOETOMTHEOOQILTFSEGMTPTSTISYTDPLDDMGYLNS
 VDDGNHSIIDFLORPINIONIEMSTNNAKTMALPLDPIINNSMYKAKERRY
 GPRADVEKLOVNAOPROAGRLILVITGYCTIGEDRKRYTDDRTNDDALVPLTSS
 PRVDDLSCTEATWCPIYSFYPSDLINVGHI GRKVVVYSPLDVSGASGLVDTG
 LMINPKNIKIKPTAMPAAQAQVTEALIQSAGAVISAASAVSSVLSPLQVPIY
 SNYARPALMVENITIDVAKHFGMSKPTTAEAPHNKLNGSGAVNAGVDMSHGLIS
 AINLEVPALPRTIDEMTVAHILARPCFIKKOMTNGSGAVNAGVLAATPTTPEFSI
 AVDSRRVASHLAYISSPFTMRGGINHFEVCTKHSGVRILTFPGDVSNDKYL
 DNADPNASXSVLRSPTDYTRFVNPVPSVOPKLSADYSRPAKDOHSGRGLYIL
 LNEBRATISVDTISCLVEVSGASDFELSMFRQKPIFTLEAKSTILNRTIRGIAV
 NVASTPVSPEIIOQTGVEGSSMKQPGSTFSSALVGGKVTSLAQILKRFLLIS
 NLTTSTNNNLYRINSFPTKPIAQAATWNIIDISYYSYIFAAYRGSFRKRIAPFN
 KYVARIILIPENGTEANAGVQEDNTVIDETRLADVMPARLNEGFEFQVHYR
 YPLPNTGASIPVIGVDLIQRNLVSNIITLEAETTKALFYRAVGDDFSFGLLGPPF
 VTRCNTILTK"

ORIGIN

Query Match 20.1%; Score 35.2; DB 13; Length 9812;
 Best Local Similarity 48.4%; Pred. No. 0.71;
 Matches 31; Conservative 15; Mismatches 18; Indels 0; Gaps 0;

QY 64 GGUUAGCUAUVUAGCUUACUAAAGACGCCGUCGCGACGCCACAAAGUCUAGUA 123
 Db 6711 GCTTAGTTATTTAACTTATCAAGATGCGCGTGGCAGCCCAAGAAATCTAGATT 6770
 QY 124 CGUC 127
 Db 6771 AGTC 6774

RESULT 8
 LOCUS AX763949 9829 bp RNA linear PAT 25-JUN-2003
 ACCESSION AX763949
 VERSION AX763949.1 GI:32258304
 KEYWORDS
 ORGANISM
 Aphid lethal paralysis virus
 Aphid lethal paralysis virus
 Viruses; ssRNA positive-strand viruses, no DNA stage;
 Diictetoviridae; Citripavinae.
 REFERENCE
 1 van der Wilk, F. and van Munster, M.

TITLE
 JOURNAL
 FEATURES
 SOURCE
 1. 9829
 /organism="Aphid lethal paralysis virus"
 /mol_type="unassigned RNA"
 /db_xref="taxon:209529"

Query Match 20.1%; Score 35.2; DB 6; Length 9829;
 Best Local Similarity 48.4%; Pred. No. 0.71;
 Matches 31; Conservative 15; Mismatches 18; Indels 0; Gaps 0;

QY 64 GGUUAGCUAUVUAGCUUACUAAAGACGCCGUCGCGACGCCACAAAGUCUAGUA 123
 Db 6711 GCTTAGTTATTTAACTTATCAAGATGCGCGTGGCAGCCCAAGAAATCTAGATT 6770
 QY 124 CGUC 127
 Db 6771 AGTC 6774

RESULT 9
 LOCUS AC107353 240387 bp DNA linear HTG 08-OCT-2002
 DEFINITION Rattus norvegicus clone CH230-47B12. *** SEQUENCING IN PROGRESS
 ACCESSION AC107353
 VERSION AC107353.4 GI:23101281
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Bkaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Murinae; Rattus.
 1 (bases 1 to 240387)
 Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooke, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, C., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., Davis, C., Davy-Carroli, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Derramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Frazer, C.M., Gabriel, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregeorgis, E., Geer, K., Giller, R., Grady, M., Guerra, M., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, J.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idetbird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kover, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenzow, L., Louie, H., Lozano, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, B.,
 Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenan, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokwelen, O., Okunolu, G., Olarnpunagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,

Nwaokemele, O., Okwomu, G., Olarnpunagoon, A., Pal, S., Parke, K.,
 Pasernak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L., U.,
 Puro, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rivers, C., Rodkey, T., Rojas, A., Rose, R., Ruiz, S. J.,
 Shender, M., Sherry, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisason, I., Sitter, C. D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorrell, R., Soza, J.,
 Steinhilber, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlarczyk, R., Wodden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstein, G. and Gibbs, R. A.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 JOURNAL

Submitted (21-Apr-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 272055)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (12-Oct-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Oct 12, 2002 this sequence version replaced gi:21747309.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

COMMENT

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GVMR
 Center clone name: CH230-248K16
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 197058 bases at least Q40
 Consensus quality: 200471 bases at least Q30
 Consensus quality: 202766 bases at least Q20
 Estimated insert size: 195853; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 3962: contig of 3962 bp in length
 * 3963 4062: gap of unknown length

FEATURES	source
4063 24797: contig of 243735 bp in length	*
24798 24798: gap of unknown length	*
24798 261437: contig of 13540 bp in length	*
261438 261537: gap of unknown length	*
261538 267069: contig of 5532 bp in length	*
267070 267169: gap of unknown length	*
267170 268553: contig of 1384 bp in length	*
268554 270058: gap of unknown length	*
270059 270158: contig of 1405 bp in length	*
270159 272055: contig of 1897 bp in length.	*
Location/Qualifiers	
1. .272055	
/organism="Rattus norvegicus"	
/mol_type="genomic DNA"	
/db_xref="taxon:10116"	
/clone="CH230-248K16"	
1. .1056	
/note="wgs_end_extension"	
clone_end:177	
3929_4782	
/note="clone_boundary"	
clone_end:177	
site:MBOT	
end_sequence:RXABF68TJ"	
3963. .4062	
/estimated_length=unknown	
4063. 5805	
/note="wgs_contig"	
225141. .226350	
/note="wgs_contig"	
complement(240781. .241693)	
/note="clone_boundary"	
clone_end:Sp6	
site:MBOT	
end_sequence:RXABF68TV"	
246372. .247797	
/note="wgs_end_extension"	
clone_end:Sp6"	
247798. .247897	
/estimated_length=unknown	
247898. .249280	
/note="wgs_end_extension"	
clone_end:Sp6"	
256950. .258196	
/note="wgs_end_extension"	
clone_end:Sp6"	
261438. .261537	
/estimated_length=unknown	
261538. .262701	
/note="wgs_end_extension"	
clone_end:Sp6"	
267070. .267169	
/estimated_length=unknown	
268554. .268653	
/estimated_length=unknown	
270059. .270158	
/estimated_length=unknown	
ORIGIN	
Query Match	20.1%; Score 35.2; DB 14; Length 272055;
Best Local Similarity	37.5%; Pred. No. 0.8; Mismatches 51; Conservative 22; Indels 0; Gaps 0;
Matches	
13 UCUCUGCGCAUUAUGUCGUAAGCGUAGCGUUAUGUCUACAACUCUGUUGACG	72
17720 TCAATTGGGATGAGAGATGATGATTAATAATCCCGGTATTAAGAAATATATTTTCGTA	17779
73 UUUAGCUUUAUCUAUACAAGACGCCGUGACGCCACCAACAAGUCUAUAUACGACAG	132
17780 TTGTGTGTAGTAAACAAGACCTTTCTTATAACACACTGATTTTAATATGTCCAGAGA	17839
133 AGAGCAUACGCGUAGGU	148

Db 17840 AAACAGCGCATGCT 17855

RESULT 11
AP008209_158/c
WPCOMMENT

Sequence split into 362 fragments LOCUS AP008209 Accession AP008209

Fragment Name	Begin	End
AP008209_000	1	110000
AP008209_001	100001	210000
AP008209_002	200001	310000
AP008209_003	300001	410000
AP008209_004	400001	510000
AP008209_005	500001	610000
AP008209_006	600001	710000
AP008209_007	700001	810000
AP008209_008	800001	910000
AP008209_009	900001	1010000
AP008209_010	1000001	1110000
AP008209_011	1100001	1210000
AP008209_012	1200001	1310000
AP008209_013	1300001	1410000
AP008209_014	1400001	1510000
AP008209_015	1500001	1610000
AP008209_016	1600001	1710000
AP008209_017	1700001	1810000
AP008209_018	1800001	1910000
AP008209_019	1900001	2010000
AP008209_020	2000001	2110000
AP008209_021	2100001	2210000
AP008209_022	2200001	2310000
AP008209_023	2300001	2410000
AP008209_024	2400001	2510000
AP008209_025	2500001	2610000
AP008209_026	2600001	2710000
AP008209_027	2700001	2810000
AP008209_028	2800001	2910000
AP008209_029	2900001	3010000
AP008209_030	3000001	3110000
AP008209_031	3100001	3210000
AP008209_032	3200001	3310000
AP008209_033	3300001	3410000
AP008209_034	3400001	3510000
AP008209_035	3500001	3610000
AP008209_036	3600001	3710000
AP008209_037	3700001	3810000
AP008209_038	3800001	3910000
AP008209_039	3900001	4010000
AP008209_040	4000001	4100000
AP008209_041	4100001	4210000
AP008209_042	4200001	4310000
AP008209_043	4300001	4410000
AP008209_044	4400001	4510000
AP008209_045	4500001	4610000
AP008209_046	4600001	4710000
AP008209_047	4700001	4810000
AP008209_048	4800001	4910000
AP008209_049	4900001	5010000
AP008209_050	5000001	5110000
AP008209_051	5100001	5210000
AP008209_052	5200001	5310000
AP008209_053	5300001	5410000
AP008209_054	5400001	5510000
AP008209_055	5500001	5610000
AP008209_056	5600001	5710000
AP008209_057	5700001	5810000
AP008209_058	5800001	5910000
AP008209_059	5900001	6010000
AP008209_060	6000001	6110000
AP008209_061	6100001	6210000
AP008209_062	6200001	6310000
AP008209_063	6300001	6410000

AP008209_064	6400001	6510000
AP008209_065	6500001	6610000
AP008209_066	6600001	6710000
AP008209_067	6700001	6810000
AP008209_068	6800001	6910000
AP008209_069	6900001	7010000
AP008209_070	7000001	7110000
AP008209_071	7100001	7210000
AP008209_072	7200001	7310000
AP008209_073	7300001	7410000
AP008209_074	7400001	7510000
AP008209_075	7500001	7610000
AP008209_076	7600001	7710000
AP008209_077	7700001	7810000
AP008209_078	7800001	7910000
AP008209_079	7900001	8010000
AP008209_080	8000001	8110000
AP008209_081	8100001	8210000
AP008209_082	8200001	8310000
AP008209_083	8300001	8410000
AP008209_084	8400001	8510000
AP008209_085	8500001	8610000
AP008209_086	8600001	8710000
AP008209_087	8700001	8810000
AP008209_088	8800001	8910000
AP008209_089	8900001	9010000
AP008209_090	9000001	9110000
AP008209_091	9100001	9210000
AP008209_092	9200001	9310000
AP008209_093	9300001	9410000
AP008209_094	9400001	9510000
AP008209_095	9500001	9610000
AP008209_096	9600001	9710000
AP008209_097	9700001	9810000
AP008209_098	9800001	9910000
AP008209_099	9900001	10010000
AP008209_100	10000001	10110000
AP008209_101	10100001	10210000
AP008209_102	10200001	10310000
AP008209_103	10300001	10410000
AP008209_104	10400001	10510000
AP008209_105	10500001	10610000
AP008209_106	10600001	10710000
AP008209_107	10700001	10810000
AP008209_108	10800001	10910000
AP008209_109	10900001	11010000
AP008209_110	11000001	11110000
AP008209_111	11100001	11210000
AP008209_112	11200001	11310000
AP008209_113	11300001	11410000
AP008209_114	11400001	11510000
AP008209_115	11500001	11610000
AP008209_116	11600001	11710000
AP008209_117	11700001	11810000
AP008209_118	11800001	11910000
AP008209_119	11900001	12010000
AP008209_120	12000001	12110000
AP008209_121	12100001	12210000
AP008209_122	12200001	12310000
AP008209_123	12300001	12410000
AP008209_124	12400001	12510000
AP008209_125	12500001	12610000
AP008209_126	12600001	12710000
AP008209_127	12700001	12810000
AP008209_128	12800001	12910000
AP008209_129	12900001	13010000
AP008209_130	13000001	13110000
AP008209_131	13100001	13210000
AP008209_132	13200001	13310000
AP008209_133	13300001	13410000
AP008209_134	13400001	13510000
AP008209_135	13500001	13610000
AP008209_136	13600001	13710000

AP008209_137	1370001	1381000
AP008209_138	1380001	1391000
AP008209_139	1390001	1401000
AP008209_140	1400001	1411000
AP008209_141	1410001	1421000
AP008209_142	1420001	1431000
AP008209_143	1430001	1441000
AP008209_144	1440001	1451000
AP008209_145	1450001	1461000
AP008209_146	1460001	1471000
AP008209_147	1470001	1481000
AP008209_148	1480001	1491000
AP008209_149	1490001	1501000
AP008209_150	1500001	1511000
AP008209_151	1510001	1521000
AP008209_152	1520001	1531000
AP008209_153	1530001	1541000
AP008209_154	1540001	1551000
AP008209_155	1550001	1561000
AP008209_156	1560001	1571000
AP008209_157	1570001	1581000
AP008209_158	1580001	1591000
AP008209_159	1590001	1601000
AP008209_160	1600001	1611000
AP008209_161	1610001	1621000
AP008209_162	1620001	1631000
AP008209_163	1630001	1641000
AP008209_164	1640001	1651000
AP008209_165	1650001	1661000
AP008209_166	1660001	1671000
AP008209_167	1670001	1681000
AP008209_168	1680001	1691000
AP008209_169	1690001	1701000
AP008209_170	1700001	1711000
AP008209_171	1710001	1721000
AP008209_172	1720001	1731000
AP008209_173	1730001	1741000
AP008209_174	1740001	1751000
AP008209_175	1750001	1761000
AP008209_176	1760001	1771000
AP008209_177	1770001	1781000
AP008209_178	1780001	1791000
AP008209_179	1790001	1801000
AP008209_180	1800001	1811000
AP008209_181	1810001	1821000
AP008209_182	1820001	1831000
AP008209_183	1830001	1841000
AP008209_184	1840001	1851000
AP008209_185	1850001	1861000
AP008209_186	1860001	1871000
AP008209_187	1870001	1881000
AP008209_188	1880001	1891000
AP008209_189	1890001	1901000
AP008209_190	1900001	1911000
AP008209_191	1910001	1921000
AP008209_192	1920001	1931000
AP008209_193	1930001	1941000
AP008209_194	1940001	1951000
AP008209_195	1950001	1961000
AP008209_196	1960001	1971000

Query Match 19.9%; Score 34.8; DB 15; Length 110000;
Best Local Similarity 39.8%; Pred. No. 1.1; Mismatches 52; Indels 0; Gaps 0;
Matches 47; Conservative 19; Mismatches 52; Indels 0; Gaps 0;

Qy	55	ACAACACUUGUACUUAUUCUUAUACAAGACGCCGUGACCCCAAAA	114
Db	73139	AAACTTGAGATGCCCATGATGCGCTTAATAATTCATTTCCCTCCCAAAA	73140
Qy	115	GUCUAGUACGUCACAGAGACAUAACGUCGUCGUCUUAUUAUAUA	172
Db	73139	AACTTAAGCAATCAATTAAGATTAATCCACATGACAAATTAATTAATTTGA	73082

RESULT 12	AC079853/c	152423 bp	DNA	linear	PLN 12-MAY-2001
LOCUS	AC079853				
DEFINITION	Oryza sativa, complete sequence.				
ACCESSION	AC079853				
VERSION	AC079853.2	GI:13384340			
KEYWORDS	HTG.				
ORGANISM	Oryza sativa				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Angiosperms; Magnoliophyta; Liliopsida; Poales; Poaceae;				
	Echinozoideae; Oryzae; Oryza.				
REFERENCE	1 (bases 1 to 152423)				
AUTHORS	Wing,R.A., Frisch,D., Presting,G., Wood,T., Yu,Y., Soderlund,C.,				
	Kim,H.-R., Rambo,T., Henry,D. and Simmons,J.				
TITLE	Rice Genomic Sequence				
JOURNAL	Unpublished				
AUTHORS	2 (bases 1 to 152423)				
	Wing,R.A., Frisch,D., Presting,G., Wood,T., Yu,Y., Rambo,T.,				
	Henry,D. and Simmons,J.				
TITLE	Direct Submision				
JOURNAL	Submitted (14-SEP-2000) Clemson University Genomics Institute,				
AUTHORS	3 (bases 1 to 152423)				
	Wing,R.A., Frisch,D., Presting,G., Wood,T., Yu,Y., Soderlund,C.,				
	Kim,H.-R., Rambo,T., Henry,D. and Simmons,J.				
TITLE	Direct Submision				
JOURNAL	Submitted (11-APR-2001) Clemson University Genomics Institute,				
AUTHORS	4 (bases 1 to 152423)				
	Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA				
REFERENCE	Wing,R.A., Frisch,D., Presting,G., Wood,T.C., Yu,Y., Soderlund,C.,				
AUTHORS	Kim,H., Rambo,T., Henry,D., Simone,J., Thurmond,S.K. and Mao,L.				
TITLE	Direct Submision				
JOURNAL	Submitted (12-MAY-2001) Clemson University Genomics Institute,				
AUTHORS	Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA				
COMMENT	On Mar 20, 2001 this sequence version replaced GI:10122028.				
	This sequence was finished as follows unless otherwise noted: all				
	regions were double stranded, sequenced with an alternate				
	chemistry, or covered by high quality data (i.e., piped quality				
	>30: an attempt was made to resolve all sequencing problems, such				
	as compressions and repeats: all regions were covered by at least				
	one plasmid subclone or more than one M13 subclone; and the				
	assembly was confirmed by restriction digest. Within the assembly				
	from 57011-57037 there are unresolved GC compressions.				
FEATURES	Location/Qualifiers				
source	1. 152423				
	/organism="Oryza sativa"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:4530"				
	/chromosome="3"				
	/clone="OSUNB0004M10"				
	155. .287				
repeat_region	/note="Similar to Stowaway01 type MITE element"				
	342. .499				
repeat_region	/note="Similar to pSIN1r6 gene, repeat sequence"				
	1664. .2077				
gene	/gene="OSUNB0004M10.1"				
	/note="Hypothetical protein"				
CDS	1664. .2077				
	/gene="OSUNB0004M10.1"				
	/codon_start=1				
	/product="Hypothetical protein"				
	/protein_id="AAK52552.1"				
	/db_xref="GI:114029011"				
	/translation="MSSSKRLNLFVIGIIFPLSDMIIVCSQGRPLIAEAAAAAQQ				
	ORHLSSSSAPRSGDVEAAAGGCGKGTAAATGCTLSPDAESGGGGGGGIVIED				
	ARPTAGHSGAGHAFTNKGVGRRLLVVTISFLI"				
	2353. .2467				
	/note="Similar to GaljIn_Os2 MITE element"				
	5369. .6848				
	/gene="OSUNB0004M10.2"				
	/note="Hypothetical protein"				


```
CDS      join(5569. .5798,6043. .6367,6792. .6848)
         /gene="OSJNB0004M10.2"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52553.1"
         /db_xref="GI:14029012"
         /translation="MTAPPMWLPGMAELRETAVAELEPHYRLRGLOAORTRRAP
SPRRDDGDEPDLRPPRSRYHELEVEGEFRHGDVAVDDASTRMTVSLMA9P
LARHNAATDRSSGGQPKTAIVPFTWSPMRLLAAVSSLDIVGSKLAPTVHG
ASSTLRSLSSQWCSRAPRRTWAPYRIFSLTSLGPEFCMG"
         complement(8290. .11889)
         /gene="OSJNB0004M10.3"
         /note="unknown protein"
         complement(join(8290. .8463,8797. .8976,10048. .10306,
10386. .10621,10933. .10998,11806. .11889))
         /gene="OSJNB0004M10.3"
         /codon_start=1
         /product="unknown protein"
         /protein_id="AAK52554.1"
         /db_xref="GI:14029013"
         /translation="MGSKAEKPKPAGGGGGEDEWCTQFGNKKKGIQFDSLACDPLIK
EYDGDALIVKSKSPPLARVVFMTMTCGIFCTMCKQIGSDLSLIVKIEVA
EOLCNKSAIILHSEVHPYQPIYTSSECKTVRFPAISQSGSGMPELINS
MNVSGEIEFSKERRNSISITKTLKQVNLDNSSAKRECTAAGLAKMLAQGLV
AMHADIADYENRRGVSAIFLRMLNLQGLVSLANNDRYKOLNTHKAVHVAYEK
LIDVLDLFLKPRSLKSRHVKIHKPISLSEQLEMDDEVYNALNGYESFLNADYRI"
         13975. .14118
         /note="Similar to Stowaway01 type MITE element"
         /protein_id="AAK52548.1"
         /db_xref="GI:14029007"
         /translation="MEVLTNDHGTARTRDGEETASVVENPDSHEMEKLTINLITE
NGCAPTHVPSLPTDVTYSLKLPQESDLSLKDPLDTSTSKTEVNNISENGSTNGST
MLSDESPKTDKNNHENTIAATNNKKEETAPRPSVYRGLIDTAPESVSEAVTKRG
GIYDKAYRSGTLERRVMQLEKAOQELPOPEQDSEACEMALQVVELEGRRRLV
ELNKKLEKRAEIDVDQKQDSELAQLQAEBOAISAKEIETREVEELTELIIKES
DELKLVKEERSTHEKHTVLASERDIAKAEBOAISAKEIETREVEELTELIIKES
LESNHAHNEAEHKLGAALAKEQDCIAMEKELOQOEELQQLMLQVSTKDAKSID
ENTHMLQILSKELAAIYENKSEBAVYIEBGSDEAKEISRSIKRALSTRKELEGV
GNTIEKADENTMLRAIETISIKSEYDKESASLVTLQOEGMASIIVSSLEALINTKEE
IEMVYIEATREKMAELPKMLQQAQAEADADAYAPHSABOELPKAKEEABOETAATAA
TAETRLRAVLEIEASFASEKLAIVAAQALOESSETSVESDPTITLPISEYHSLSK
RYVEAELEAERVAALAAQIELAEASSTRLEIRQOETKEHKKKDALQIALQADRA
KEGILGAEOELRKRAELIQRKAKGVANPWTAPPRSPBOKSKYEDDAVLTEPNS
PMSNSTDDPVVDQIKRKKTFPQMSIISRKAQT"
         43860. .44040
         /note="Similar to Tourist03 type MITE element"
         /codon_start=1
         /product="Putative heme oxygenase 2"
         /protein_id="AAK52555.1"
         /db_xref="GI:14029014"
         /translation="MPLAAVAVASAVVPPRPPRRARPLRSFTGLILTRDLAALT
VAKCAPSPRAAABEAABAIVADVDAAPKPRPRPRPYRQYRGEAGVAEEMFVAMR
NPKRTLLKMDTGAEEVDGVSIEDASSEEEEDDDVVEEERGAGLEGMMNS
MEGFVKYLVDSKLVDETVERTVASTDVAAYVFRKSGERSARITKDLQEWGGLAV
PEPSTAGSTVATVTELASNAPAFSLSHYNIYPAHTTGVAIGNKISKILIEGREL
FYKDSVVELLKDTREKLNELSLKVSPELDISICRRRLPHKRGDEWVSRMLIGPRHY
GLALPIAMAAEAAMGCP"
         27869. .28076
         /note="Similar to Wanderer_011 MITE element"
         /protein_id="AAK52556.1"
         /db_xref="GI:14029015"
         /translation="MPLAAVAVASAVVPPRPPRRARPLRSFTGLILTRDLAALT
VAKCAPSPRAAABEAABAIVADVDAAPKPRPRPRPYRQYRGEAGVAEEMFVAMR
NPKRTLLKMDTGAEEVDGVSIEDASSEEEEDDDVVEEERGAGLEGMMNS
MEGFVKYLVDSKLVDETVERTVASTDVAAYVFRKSGERSARITKDLQEWGGLAV
PEPSTAGSTVATVTELASNAPAFSLSHYNIYPAHTTGVAIGNKISKILIEGREL
FYKDSVVELLKDTREKLNELSLKVSPELDISICRRRLPHKRGDEWVSRMLIGPRHY
GLALPIAMAAEAAMGCP"
         30319. .30494
         /note="Similar to Wanderer_Os1 MITE element"
         /protein_id="AAK52557.1"
         /db_xref="GI:14029016"
         /translation="MPLAAVAVASAVVPPRPPRRARPLRSFTGLILTRDLAALT
VAKCAPSPRAAABEAABAIVADVDAAPKPRPRPRPYRQYRGEAGVAEEMFVAMR
NPKRTLLKMDTGAEEVDGVSIEDASSEEEEDDDVVEEERGAGLEGMMNS
MEGFVKYLVDSKLVDETVERTVASTDVAAYVFRKSGERSARITKDLQEWGGLAV
PEPSTAGSTVATVTELASNAPAFSLSHYNIYPAHTTGVAIGNKISKILIEGREL
FYKDSVVELLKDTREKLNELSLKVSPELDISICRRRLPHKRGDEWVSRMLIGPRHY
GLALPIAMAAEAAMGCP"
         30627. .31005
         /note="Similar to Crackle01 MITE element"
         /protein_id="AAK52558.1"
         /db_xref="GI:14029017"
         /translation="MPLAAVAVASAVVPPRPPRRARPLRSFTGLILTRDLAALT
VAKCAPSPRAAABEAABAIVADVDAAPKPRPRPRPYRQYRGEAGVAEEMFVAMR
NPKRTLLKMDTGAEEVDGVSIEDASSEEEEDDDVVEEERGAGLEGMMNS
MEGFVKYLVDSKLVDETVERTVASTDVAAYVFRKSGERSARITKDLQEWGGLAV
PEPSTAGSTVATVTELASNAPAFSLSHYNIYPAHTTGVAIGNKISKILIEGREL
FYKDSVVELLKDTREKLNELSLKVSPELDISICRRRLPHKRGDEWVSRMLIGPRHY
GLALPIAMAAEAAMGCP"
         31743. .31879
         /note="Similar to Stowaway01 type MITE element"
         /protein_id="AAK52559.1"
         /db_xref="GI:14029018"
         /translation="MPLAAVAVASAVVPPRPPRRARPLRSFTGLILTRDLAALT
VAKCAPSPRAAABEAABAIVADVDAAPKPRPRPRPYRQYRGEAGVAEEMFVAMR
NPKRTLLKMDTGAEEVDGVSIEDASSEEEEDDDVVEEERGAGLEGMMNS
MEGFVKYLVDSKLVDETVERTVASTDVAAYVFRKSGERSARITKDLQEWGGLAV
PEPSTAGSTVATVTELASNAPAFSLSHYNIYPAHTTGVAIGNKISKILIEGREL
FYKDSVVELLKDTREKLNELSLKVSPELDISICRRRLPHKRGDEWVSRMLIGPRHY
GLALPIAMAAEAAMGCP"
         32992. .36053
         /gene="OSJNB0004M10.5"
         /note="unknown protein"
         complement(join(32992. .33231,33969. .34140,34322. .34475,34553. .34616,
34799. .34906,35083. .35208,35503. .35535,35831. .36053))
         /gene="OSJNB0004M10.5"
         /codon_start=1
         /product="unknown protein"
         /protein_id="AAK52556.1"
         /db_xref="GI:14029015"
```

```
/translation="MEVGEVDGSGVPLAVLKLKRELGNQKVERPDMLFGEASKSKEE
DPTFLPKSGRGOAODEDAGGADDDTISVFAIFDGNNGAALITRENTLNNY
LAIPRUTSREBETTLPRALVNGRVUTDEPOTKARSGCTGTFVITDGVTVAVS
GDSRCTIESAGSVYFLSADHRLDTNEEBERTVYASGDVGRINISAGAGIGPRCP
GGICLSRISIDIDGERTVPPVHKQVLSNAGARVIADGDALRFOEALVYTG
LPAAASRIVSKESGSKGRDPTCVIVLIPPEKLSPLKKGKGGIKALFFRRS
DELTEQDMDGCLPEPDVVEIEYEGSSMLAKORLKTINPTNMFMLHDCACVOLVMBG
EGISVGSITRNSVDPDWPGLCSGQLKKEAMKQHLTSMNAF"
         complement(37031. .37604)
         /gene="OSJNB0004M10.6"
         /note="Putative calmodulin-like protein"
         complement(join(37031. .37053,37086. .37216,37291. .37604))
         /gene="OSJNB0004M10.6"
         /codon_start=1
         /product="Putative calmodulin-like protein"
         /protein_id="AAK52557.1"
         /db_xref="GI:14029016"
         /translation="MVLPSITLKI PAMCRALISSPPMTLEBKFMKENTNDR
ISGVELEAIRSKGFGSAMWKSIVLHQADKDRNGYIDFEIENLVTPAQKVGIKI
TTMOQHLDNOKAVKGLAVTSVSLAVHFDNNEQLKLNPFWFAALFESI"
         complement(39422. .41977)
         /gene="OSJNB0004M10.7"
         /note="Putative myosin heavy chain-like"
         complement(join(39422. .41230,41456. .41977))
         /gene="OSJNB0004M10.7"
         /codon_start=1
         /product="Putative myosin heavy chain-like"
         /protein_id="AAK52548.1"
         /db_xref="GI:14029007"
         /translation="MEVLTNDHGTARTRDGEETASVVENPDSHEMEKLTINLITE
NGCAPTHVPSLPTDVTYSLKLPQESDLSLKDPLDTSTSKTEVNNISENGSTNGST
MLSDESPKTDKNNHENTIAATNNKKEETAPRPSVYRGLIDTAPESVSEAVTKRG
GIYDKAYRSGTLERRVMQLEKAOQELPOPEQDSEACEMALQVVELEGRRRLV
ELNKKLEKRAEIDVDQKQDSELAQLQAEBOAISAKEIETREVEELTELIIKES
DELKLVKEERSTHEKHTVLASERDIAKAEBOAISAKEIETREVEELTELIIKES
LESNHAHNEAEHKLGAALAKEQDCIAMEKELOQOEELQQLMLQVSTKDAKSID
ENTHMLQILSKELAAIYENKSEBAVYIEBGSDEAKEISRSIKRALSTRKELEGV
GNTIEKADENTMLRAIETISIKSEYDKESASLVTLQOEGMASIIVSSLEALINTKEE
IEMVYIEATREKMAELPKMLQQAQAEADADAYAPHSABOELPKAKEEABOETAATAA
TAETRLRAVLEIEASFASEKLAIVAAQALOESSETSVESDPTITLPISEYHSLSK
RYVEAELEAERVAALAAQIELAEASSTRLEIRQOETKEHKKKDALQIALQADRA
KEGILGAEOELRKRAELIQRKAKGVANPWTAPPRSPBOKSKYEDDAVLTEPNS
PMSNSTDDPVVDQIKRKKTFPQMSIISRKAQT"
         43860. .44040
         /note="Similar to Tourist03 type MITE element"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene="OSJNB0004M10.8"
         /codon_start=1
         /product="Hypothetical protein"
         /protein_id="AAK52549.1"
         /db_xref="GI:14029008"
         /translation="MAERDRKTRGVAEGRDRDRSGTMDGAPPARGERGSAVAGPL
TEGSRBAWGTADVEGRDDREYVGEHGGDCGVPAGGROUGARGAATFVKQGEAA
GLILGAAMSSRRQAGADQVGGVGEAGIGFVGLVALGDSLSRN"
         50104. .50877
         /gene="OSJNB0004M10.8"
         /note="Hypothetical protein"
         join(50104. .50332,50651. .50877)
         /gene
```


This Page Blank (usp10)

CC viruses (AB280707-AB280713). This sequence represents the "pseudoknot"
CC higher-order sequence from the Rhopalosiphum virus. The sequence is used
CC in a construct which may also include an intergenic region and internal
CC ribosome entry site (IGR-IRIS). The method is applicable in producing
CC useful proteins
CC
CC

50 Sequence 175 BP; 47 A; 38 C; 40 G; 0 T; 50 U; 0 Other;

Query Match 100.0%; Score 175; DB 8; Length 175;
Best Local Similarity 100.0%; Pred. No. 1.6e-53;

Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGUGUGUGUAGUACUUGCGCGAUAUAAGUGUAGCGUAAAACGUGGCUAUAUAGCA 60
Db 1 AGUGUGUGUAGUACUUGCGCGAUAUAAGUGUAGCGUAAAACGUGGCUAUAUAGCA 60

Qy 61 CUUGGUUAGCUAUAUAGCUUUAACUUAUCAAAGACGCCGUGUGCAAGCCCAAAAGUCUAG 120
Db 61 CUUGGUUAGCUAUAUAGCUUUAACUUAUCAAAGACGCCGUGUGCAAGCCCAAAAGUCUAG 120

Qy 121 AUAAGUCACAGAGAGCAUAACGUAGUGCGGUGUACUAUCCUUAUAUAGACCU 175
Db 121 AUAAGUCACAGAGAGCAUAACGUAGUGCGGUGUACUAUCCUUAUAUAGACCU 175

RESULT 2

1D AAL50550 standard; RNA; 187 BP.

XX AAL50550;

DT 19-DEC-2002 (first entry)

DE Rhopalosiphum padi virus RNA sequence.

XX CrPV-like virus; ss, higher-order structure; drug development;
KW drug production; translational activity-promoting function;
KW protein synthesis; structural analysis.

XX Rhopalosiphum padi virus.

XX WO200261080-A1.

PD 08-AUG-2002.

PF 31-JAN-2001; 2001WO-JP000641.

PR 25-JAN-2001; 2001JP-00016746.

PA (NNAAG-) NAT INST AGROBIOLOGICAL SCI.

PI Nakashima N, Kanamori Y;

DR WPI; 2002-627482/67.

XX Translational activity-promoting higher-order structure of CrPV-like
XX viruses for protein translation when suitably initiated, useful in
XX synthesis of proteins and polypeptides of foreign species for application
XX in drugs.

PS Claim 1; Fig 1-2; 38pp; Japanese.

XX The invention comprises seven RNA sequences (CrPV-like viruses) which
XX have a higher-order structure that sustains translational activity-
XX promoting function. The RNA sequences of the invention are useful in the
XX synthesis of proteins and polypeptides for application in developing and
XX producing drugs. The RNA sequences of the invention are also useful in
XX basic research of protein synthesis and structural analysis by the gene
XX recombinant technique. The present nucleotide sequence represents a Rhopalosiphum
XX padi virus RNA sequence of the invention

50 Sequence 187 BP; 54 A; 40 C; 41 G; 0 T; 52 U; 0 Other;

Query Match 100.0%; Score 175; DB 6; Length 187;
Best Local Similarity 100.0%; Pred. No. 1.6e-53;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGUGUGUGUAGUACUUGCGCGAUAUAAGUGUAGCGUAAAACGUGGCUAUAUAGCA 60
Db 1 AGUGUGUGUAGUACUUGCGCGAUAUAAGUGUAGCGUAAAACGUGGCUAUAUAGCA 60

Qy 61 CUUGGUUAGCUAUAUAGCUUUAACUUAUCAAAGACGCCGUGUGCAAGCCCAAAAGUCUAG 120
Db 61 CUUGGUUAGCUAUAUAGCUUUAACUUAUCAAAGACGCCGUGUGCAAGCCCAAAAGUCUAG 120

Qy 121 AUAAGUCACAGAGAGCAUAACGUAGUGCGGUGUACUAUCCUUAUAUAGACCU 175
Db 121 AUAAGUCACAGAGAGCAUAACGUAGUGCGGUGUACUAUCCUUAUAUAGACCU 175

RESULT 3

1D AAL60324 standard; RNA; 9829 BP.

XX AAL60324;

DT 27-AUG-2003 (first entry)

DE Aphid lethal paralysis virus (ALPV) RNA.

XX ALPV; Aphid lethal paralysis virus; homopterous insect; floriculture;
KW aphid; white fly; agriculture; gene therapy; ss.

XX Aphid lethal paralysis virus.

XX WO2003040177-A2.

PD 15-MAY-2003.

PF 25-OCT-2002; 2002WO-NL000682.

PR 25-OCT-2001; 2001NL-01019225.

PA (PLAN-) PLANT RES INT BV.

PI Van Der Wilk F, Van Munster M;

DR WPI; 2003-441523/41.

XX Novel nucleic acid derived from aphid lethal paralysis virus, useful for
XX detecting ALPV-viruses, for preparing a monoclonal antibody against
XX antigen, in research concerning aphids and for controlling aphids.

PS Claim 1; Page 10-13; 17pp; English.

XX The invention relates to a novel RNA derived from an ALPV (Aphid lethal
XX paralysis virus). ALPV RNA is useful for controlling homopterous insects
XX such as aphids or white flies. It is useful in the research concerning
XX aphids and for controlling aphids and in the field of agriculture or
XX floriculture. It is also useful for detecting ALPV- viruses using PCR and
XX for preparing a monoclonal antibody against antigen which is synthesised
XX in vitro. The invention is also useful in gene therapy. The present
XX sequence is ALPV RNA

50 Sequence 9829 BP; 3083 A; 1906 C; 1879 G; 0 T; 2961 U; 0 Other;

Query Match 20.1%; Score 35.2; DB 8; Length 9829;
Best Local Similarity 71.9%; Pred. No. 0.095;

Matches 46; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 64 GGUUAGCUUUAUAGCUUUAACUUAUCAAAGACGCCGUGUGCAAGCCCAAAAGUCUAG 123
Db 6711 GCUUAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 6770

Qy 124 CGUC 127

Db 6771 AGUC 6774

RESULT 4

ID ADR85781 standard; DNA; 5355 BP.

XX ADR85781;

XX 04-NOV-2004 (first entry)

XX Aspergillus fumigatus essential gene open reading frame #418.

XX Fungicide; Aspergillus fumigatus infection; Farmer's lung disease;

XX drug screening; ds.

XX Aspergillus fumigatus.

XX WO2004067709-A2.

XX 12-AUG-2004.

XX 16-JAN-2004; 2004WO-US001099.

XX 17-JAN-2003; 2003US-0441281P.

XX 13-JUN-2003; 2003US-0478196P.

XX (ELIT-) ELITRA PHARM INC.

XX (ELIT-) ELITRA CANADA LTD.

XX Jiang B, Hu W, Lemieux S, Roemer T;

XX WPI; 2004-594200/57.

XX P-PSDB; ADR86368.

PT New purified or isolated Aspergillus fumigatus nucleic acid molecule
PT encoding a gene product, useful for diagnosing and/or treating invasive
PT fungal infections, such as Farmer's lung disease.

XX Claim 2; SEQ ID NO 2418; 164pp; English.

XX The present invention relates to Aspergillus fumigatus genes that are
XX essential and are potential targets for drug screening. The methods and
XX compositions of the present invention are useful for diagnosing and/or
XX treating invasive Aspergillus fumigatus infection, including the allergic
XX forms of the disease, such as Farmer's lung disease. They can also be
XX used in various drug discovery purposes, such as expression of the
XX recombinant protein, hybridization assay and construction of nucleic acid
XX arrays. The present sequence represents an Aspergillus fumigatus
XX essential gene open reading frame, used during diagnosis and drug
XX development in the invention. These genes share a high degree of sequence
XX conservation with known essential genes of candida albicans. The sequence
XX data for this patent is not represented in the printed specification, but
XX was obtained in electronic format from WIPO.

SQ Sequence 5355 BP; 1369 A; 1314 C; 1377 G; 1295 T; 0 U; 0 Other;

Query Match 19.2%; Score 33.6; DB 13; Length 5355;

Best Local Similarity 38.9%; Pred. No. 0.29;

Matches 56; Conservative 19; Mismatches 69; Indels 0; Gaps 0;

OY 3 UGUGUGUAGUCCGGAUAAUGCGUGAAGAAACGUGGUAUUGCUACAACU 62

DB 2949 TCTTCTGAGGGCGAAGCCGATTAATGGCCGATGTCAACACATTGGAACACCGAACCACT 3008

OY 63 UGUUAGCUAUVUAGCUUUAUUAUACAAGAGCCGCGUGGACGCCACAAAGUUCUAU 122

DB 3009 TCCTTACAGCTTACCATTTCTGAGCTGGAAGATGGAATGCAACAATTAATAATTCGGGG 3068

OY 123 ACUGACAGAGAGCAUACGCUAG 146

DB 3069 AAGTAGCAGAGAGACCAACCAAG 3092

RESULT 5

ID ADR85194 standard; DNA; 5442 BP.

XX ADR85194;

XX 04-NOV-2004 (first entry)

XX Aspergillus fumigatus essential gene with introns #418.

XX Fungicide; Aspergillus fumigatus infection; Farmer's lung disease;

XX drug screening; ds.

XX Aspergillus fumigatus.

XX WO2004067709-A2.

XX 12-AUG-2004.

XX 16-JAN-2004; 2004WO-US001099.

XX 17-JAN-2003; 2003US-0441281P.

XX 13-JUN-2003; 2003US-0478196P.

XX (ELIT-) ELITRA PHARM INC.

XX (ELIT-) ELITRA CANADA LTD.

XX Jiang B, Hu W, Lemieux S, Roemer T;

XX WPI; 2004-594200/57.

XX P-PSDB; ADR86368.

PT New purified or isolated Aspergillus fumigatus nucleic acid molecule
PT encoding a gene product, useful for diagnosing and/or treating invasive
PT fungal infections, such as Farmer's lung disease.

XX Claim 3; SEQ ID NO 1418; 164pp; English.

XX The present invention relates to Aspergillus fumigatus genes that are
XX essential and are potential targets for drug screening. The methods and
XX compositions of the present invention are useful for diagnosing and/or
XX treating invasive Aspergillus fumigatus infection, including the allergic
XX forms of the disease, such as Farmer's lung disease. They can also be
XX used in various drug discovery purposes, such as expression of the
XX recombinant protein, hybridization assay and construction of nucleic acid
XX arrays. The present sequence represents an Aspergillus fumigatus
XX essential gene sequence containing all introns, used during diagnosis and
XX drug development in the invention. These genes share a high degree of
XX sequence conservation with known essential genes of candida albicans. The
XX sequence data for this patent is not represented in the printed
XX specification, but was obtained in electronic format from WIPO.

SQ Sequence 5442 BP; 1392 A; 1331 C; 1397 G; 1322 T; 0 U; 0 Other;

Query Match 19.2%; Score 33.6; DB 13; Length 5442;

Best Local Similarity 38.9%; Pred. No. 0.29; Mismatches 69; Indels 0; Gaps 0;

Matches 56; Conservative 19; Mismatches 69; Indels 0; Gaps 0;

OY 3 UGUGUGUAGUCCGGAUAAUGCGUGAAGAAACGUGGUAUUGCUACAACU 62

DB 2949 TCTTCTGAGGGCGAAGCCGATTAATGGCCGATGTCAACACATTGGAACACCGAACCACT 3008

OY 63 UGUUAGCUAUVUAGCUUUAUUAUACAAGAGCCGCGUGGACGCCACAAAGUUCUAU 122

DB 3009 TCCTTACAGCTTACCATTTCTGAGCTGGAAGATGGAATGCAACAATTAATAATTCGGGG 3068

OY 123 ACUGACAGAGAGCAUACGCUAG 146

DB 3069 AAGTAGCAGAGAGACCAACCAAG 3092

RESULT 6


```
AD84607
ID AD84607 standard; DNA; 7442 BP.
XX
XX
AC AD84607;
XX
XX 04-NOV-2004 (first entry)
XX
XX Aspergillus fumigatus essential gene genomic sequence #418.
XX
XX Fungicide; Aspergillus fumigatus infection; Farmer's lung disease;
XX drug screening; ds.
XX
XX OS Aspergillus fumigatus.
XX
XX WO2004067709-A2.
XX
XX 12-AUG-2004.
XX
XX 16-JAN-2004; 2004WO-US001099.
XX
XX 17-JAN-2003; 2003US-0441281P.
XX
XX 13-JUN-2003; 2003US-0478196P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX (ELIT-) ELITRA CANADA LTD.
XX
XX Jiang B, Hu W, Lemieux S, Roemer T;
XX
XX WPI; 2004-594200/57.
XX
XX P-PSDB; AD86368.
XX
XX New purified or isolated Aspergillus fumigatus nucleic acid molecule
XX encoding a gene product, useful for diagnosing and/or treating invasive
XX fungal infections, such as Farmer's lung disease.
XX
XX Claim 3; SEQ ID NO 418; 164pp; English.
XX
XX The present invention relates to Aspergillus fumigatus genes that are
XX essential and are potential targets for drug screening. The method and
XX compositions of the present invention are useful for diagnosing and/or
XX treating invasive Aspergillus fumigatus infection, including the allergic
XX forms of the disease, such as Farmer's lung disease. They can also be
XX used in various drug discovery purposes, such as expression of the
XX recombinant protein, hybridization assay and construction of nucleic acid
XX arrays. The present sequence represents an Aspergillus fumigatus
XX essential gene full length genomic sequence, used during diagnosis and
XX drug development in the invention. These genes share a high degree of
XX sequence conservation with known essential genes of candida albicans. The
XX sequence data for this patent is not represented in the printed
XX specification, but was obtained in electronic format from WIPO.
XX
XX Sequence 7442 BP; 1925 A; 1827 C; 1878 G; 1812 T; 0 U; 0 Other;
SQ
Query Match 19.2%; Score 33.6; DB 13; Length 7442;
Best Local Similarity 38.9%; Pred. No. 0.33;
Matches 56; Conservative 19; Mismatches 69; Indels 0; Gaps 0;
OY 3 UGUGUGAUCUUGCCGGAUAAUGCUGAGUGAAAACUUGCUAUGCACAACU 62
DB 3949 TCTTCTGAGGCGACGACCGATATATGCGCATGTCTCACACATTGGAAACACGAACTCT
OY 63 UGUUAGCUAUVUAGCUUACUAAUACAAGACGCGUGUGAGCCCAACAAAGUCUAGAU 122
DB 4009 TCCTCTACAGTCTACCATTTCTGAGCTGAAAGATGGAATGCAACAACTTAAATTCGGGG 4068
OY 123 ACUCACAGAGAGACUACGCUAG 146
DB 4069 AAGTAGCAGAGAGACCAACCAAG 4092
RESULT 7
ID ABA03158 standard; DNA; 5853 BP.
```

```
XX
XX ABA03158;
XX
XX 08-FEB-2002 (first entry)
XX
XX Streptococcus thermophilus pgm gene.
XX
XX Lactic acid bacterium; exopolysaccharide; pgm; phosphoglucosyltransferase;
XX fermented food product; food additive; ds.
XX
XX OS Streptococcus thermophilus.
XX
XX Key Location/Qualifiers
XX CDS 2169..3092
XX /tag= a
XX /product= "PGM"
XX
XX WO200179476-A1.
XX
XX 25-OCT-2001.
XX
XX 18-APR-2001; 2001WO-FR001198.
XX
XX 18-APR-2000; 2000FR-00004971.
XX
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX (DANO-) CIE DANONE SA GERVAIS.
XX (RHOD ) RHODIA CHIM.
XX
XX Amba J, Hagege J, Chailion S, Beesancon-Yoshpe I, Fremaux C;
XX Mengaud J, Renault P;
XX
XX WPI; 2002-026026/03.
XX
XX P-PSDB; AAM47449.
XX
XX New mutant lactic acid bacteria, useful e.g. for making fermented foods,
XX have the pgm gene inactivated and overexpress exopolysaccharides.
XX
XX Example 2; Page 25-29; 42pp; French.
XX
XX The present invention relates to a mutant lactic acid bacterium that
XX overexpresses exopolysaccharides. The mutant bacterium has the pgm gene
XX (alpha-phosphoglucosyltransferase) at least partly inactivated, which results in
XX increased production of exopolysaccharides. The mutant bacterium can be
XX used to produce: (i) fermented food products; or (ii) exopolysaccharides,
XX which are useful as food additives and in cosmetics and pharmaceuticals,
XX e.g. as thickeners, gelling agents, texture stabilizers and fat
XX substitutes. The present sequence is the pgm gene from Streptococcus
XX thermophilus, which was used in an example from the present invention
XX
XX Sequence 5853 BP; 1646 A; 1028 C; 1342 G; 1837 T; 0 U; 0 Other;
SQ
Query Match 18.1%; Score 31.6; DB 6; Length 5853;
Best Local Similarity 40.7%; Pred. No. 1.6;
Matches 48; Conservative 16; Mismatches 54; Indels 0; Gaps 0;
OY 6 UGUGUGAUCUUGCCGGAUAAUGCUGAGUGAAAACUUGCUAUGCACAACUUG 65
DB 437 TGTCTCAAGTCCCAACGAAGTTCTGCGGCAAAATATACACTCTGCAATGACCACTA 378
OY 66 UUAACUUVUAGCUUUAUACAAGACGCGUGUGAGCCCAACAAAGUCUAGAU 123
DB 377 ATTGATGACTTGATTCTATCAATTAAGAAATCGTGTGAATTCATTAATCAGAGATA 320
RESULT 8
ID AAS72193 standard; cDNA; 2398 BP.
XX
XX AAS72193;
XX
XX 13-FEB-2002 (first entry)
XX
```


DE	DNA encoding novel human diagnostic protein #7997.
KM	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM	food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US008631.
XX	
PR	31-MAR-2000; 2000US-00540217.
XX	
PR	23-AUG-2000; 2000US-00649167.
XX	
PA	(HYSE-) HYSEQ INC.
PI	Drmanc RT, Liu C, Tang YT;
PI	
XX	WPI; 2001-639362/73.
DR	P-PSDB; ABG08006.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
PS	Claim 1; SEQ ID NO 7997; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polynucleotide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC	coding sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WFO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 2398 BP; 561 A; 563 C; 611 G; 663 T; 0 U; 0 Other;
XX	
Query Match	17.9%; Score 31.4; DB 5; Length 2398;
Best Local Similarity	37.2%; Pred. No. 1.4;
Matches	54; Conservative 20; Mismatches 71; Indels 0; Gaps 0;
OY	1 AGUGUGUGAUCUCUGCGCAUAUAUUCGUGAAAAGUGUGUAGUACUACA 60
DB	1674 AATGTGTCGGAAATTTGTGGGGAAATGAGACCAGAAAGCTCTGMAAAGTGGAA 1615
OY	61 CUGUGUAGUUAUUGUAGUUAUAUAAGACCGGUGUGUGCCGACCAAAAGUCUAG 120
DB	1614 GATGTGAGAGACATTTCAAGCAAGTATATAGAACAATTCATGCAACCAAGGAGG 1555
OY	121 AUAAGUCACAGAGAGAGCAUAAGCUA 145
DB	1554 GTCTGTACTTGGGAGATTTCGCTTA 1530
RESULT 9	
ABZ80709	

1D	ABZ80709 standard; RNA; 189 BP.	
AC	ABZ80709;	
XX		
DT	15-OCT-2003 (first entry)	
XX		
DE	Drosophila C virus derived pseudoknot sequence.	
XX		
KW	Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;	
KW	albumen; impurity; higher-order structure; intergenic region; IGR-IRES;	
KW	internal ribosome entry site; ss.	
OS	Drosophila C virus.	
XX		
PN	WO200303719-A1.	
PD	24-APR-2003.	
XX		
PF	08-OCT-2002; 2002WO-JP010447.	
XX		
PR	17-OCT-2001; 2001JP-00319923.	
XX		
PA	(NAG-) NAT INST AGROBIOLOGICAL SCI.	
XX		
PA	(WAKE-) WAKENYAKU KK.	
XX		
PI	Nakashima N, Shibuya N, Nishikawa S;	
DR	WPI; 2003-403230/38.	
XX		
PT	Cell-free protein synthesis means in wheatgerm system to establish	
PT	overexpression of target gene with base sequence sustaining translation	
XX	activity and function promotion, for producing useful proteins.	
PS	Claim 1; Page 32; 39pp; Japanese.	
XX		
CC	The invention relates to a cell-free protein synthesis system derived	
CC	from wheatgerm where there is substantial exclusion of wheatgerm embryo	
CC	albumen impurities. The novel system uses a sequence having a higher-	
CC	order RNA structure that promotes translation activity. The higher-order	
CC	sequence is preferably a "pseudoknot", especially derived from a range of	
CC	viruses (ABZ80707-ABZ80713). This sequence represents the "pseudoknot"	
CC	higher-order sequence from the Drosophila C virus. The sequence is used	
CC	in a construct which may also include an intergenic region and internal	
CC	ribosome entry site (IGR-IRES). The method is applicable in producing	
XX	useful proteins	
XX		
SO	Sequence 189 BP; 57 A; 30 C; 34 G; 0 T; 68 U; 0 Other;	
XX		
Query Match	17.8%; Score 31.2; DB 8; Length 189;	
Best Local Similarity	66.2%; Pred. No. 0.67;	
Matches	45; Conservative 0; Mismatches 23; Indels 0; Gaps 0;	
OY	53 CUACAACUUGUGUACCUAUAUAGCUUAUACAAGCGCCGUGUCAGCCCA 112	
DB	65 CUUAUAUAUUAAGGUACAUAUAUAUUAUCUGUUCAGAUAGCCUAUUGGACGCCCA 124	
OY	113 AAGUCUAG 120	
DB	125 AUUUCGAG 132	
RESULT 10		
AL50546		
ID	AL50546 standard; RNA; 201 BP.	
XX		
AC	AL50546;	
XX		
DT	19-DEC-2002 (first entry)	
XX		
DE	Drosophila C virus RNA sequence.	
XX		
KW	CirV-like virus; ss; higher-order structure; drug development;	
KW	drug production; translational activity-promoting function;	

KM		protein synthesis; structural analysis.
XX		
OS	Drosophila C virus.	
XX		
PN	WO200261080-A1.	
XX		
PD	08-AUG-2002.	
XX		
PP	31-JAN-2001; 2001WO-JP0006641.	
XX		
PR	25-JAN-2001; 2001JP-00016746.	
XX		
PA	(NAG-) NAT INST AGROBIOLOGICAL SCI.	
XX		
PI	Nakashima N, Kanamori Y;	
XX		
DR	WPI; 2002-627482/67.	
XX		
PT	Translational activity-Promoting higher-order structure of CrPV-like	
PT	viruses for protein translation when suitably initiated, useful in	
PT	synthesis of proteins and polypeptides of foreign species for application	
PT	in drugs.	
XX		
PS	Claim 1; Fig 1-2; 38pp; Japanese.	
CC		
CC	The invention comprises seven RNA sequences (CrPV-like viruses) which	
CC	have a higher-order structure that sustains translational activity-	
CC	promoting function. The RNA sequences of the invention are useful in the	
CC	synthesis of proteins and polypeptides for application in developing and	
CC	producing drugs. The RNA sequences of the invention are also useful in	
CC	basic research of protein synthesis and structural analysis by the gene	
CC	recombinant technique. The present nucleotide represents a Drosophila C	
CC	virus RNA sequence of the invention	
XX		
SQ	Sequence 201 BP; 61 A; 33 C; 35 G; 0 T; 72 U; 0 Other;	
	Query Match 17.8%; Score 31.2; DB 6; Length 201;	
	Best Local Similarity 66.2%; Pred. No. 0.64; Mismatches 23; Indels 0; Gaps 0	
	Matches 45; Conservative 0;	
OY	53 CUACAACUUGUUAGCUUAUCUAUACAAGCCGUGUGCAGCCCAAA 112	
DB	65 CUUAUAUAUAGGUAAACUAUUAAGUUUACUGUUCAGAUGCCUAUUGGACGCCCAA 124	
OY	113 AAUGCUG 120	
DB	125 AUAUCCAG 132	
	RESULT 11	
	ACA21093	
ID	ACA21093 standard; DNA; 1992 BP.	
XX		
AC	ACA21093;	
XX		
DT	19-JUN-2003 (first entry)	
DE	Prokaryotic essential gene #2750.	
XX		
KW	Antisense; ds; prokaryotic essential gene; cell proliferation;	
XX	drug design; gene.	
OS	Acinetobacter baumannii.	
XX		
PN	WO200277183-A2.	
XX		
PD	03-OCT-2002.	
XX		
PF	21-MAR-2002; 2002WO-US009107.	
XX		
PR	21-MAR-2001; 2001US-00815242.	
PR	06-SEP-2001; 2001US-00948993.	
PR	25-OCT-2001; 2001US-0342923P.	

PR		08-FEB-2002; 2002US-00072851.
PR		06-MAR-2002; 2002US-0362699P.
XX		
PA	(EUIT-) ELITRA PHARM INC.	
XX		
P1	Mang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,	
P1	Wall D, Trawick JD, Carr GD, Yamamoto R, Foreyth RA, Xu HH,	
XX		
DR	WPI: 2003-029926/02.	
XX	P-SDB: ABU17223.	
PT	New antisense nucleic acids, useful for identifying proteins or screening	
PT	for homologous nucleic acids required for cellular proliferation to	
PT	isolate candidate molecules for rational drug discovery programs.	
XX		
PS	Claim 14; SEQ ID NO 8963; 1766pp; English.	
XX		
CC	The invention relates to an isolated nucleic acid comprising any one of	
CC	the 6213 antisense sequences given in the specification where expression	
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:	
CC	(1) a vector comprising a promoter operably linked to the nucleic acid	
CC	encoding a polypeptide whose expression is inhibited by the antisense	
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated	
CC	polypeptide or its fragment whose expression is inhibited by the	
CC	antisense nucleic acid; (4) an antibody capable of specifically binding	
CC	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular	
CC	proliferation or the activity of a gene in an operon required for	
CC	proliferation; (7) identifying a compound that influences the activity of	
CC	the gene product or that has an activity against a biological pathway	
CC	required for proliferation, or that inhibits cellular proliferation; (8)	
CC	identifying a gene required for cellular proliferation or the biological	
CC	pathway in which a proliferation-required gene or its gene product lies	
CC	or a gene on which the test compound that inhibits proliferation of an	
CC	organism acts; (9) manufacturing an antibiotic; (10) profiling a	
CC	compound's activity; (11) a culture comprising strains in which the gene	
CC	product is overexpressed or underexpressed; (12) determining the extent	
CC	to which each of the strains is present in a culture or collection of	
CC	strains; or (13) identifying the target of a compound that inhibits the	
CC	proliferation of an organism. The antisense nucleic acids are useful for	
CC	identifying proteins or screening for homologous nucleic acids required	
CC	for cellular proliferation to isolate candidate molecules for rational	
CC	drug discovery programs, or for screening homologous nucleic acids	
CC	required for proliferation in cells other than S. aureus, S. typhimurium,	
CC	K. pneumoniae or P. aeruginosa. The present sequence data for this patent did	
CC	not form part of the printed specification, but was obtained in	
CC	electronic format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences	
XX		
SQ	Sequence 1992 BP; 639 A; 402 C; 410 G; 541 T; 0 U; 0 Other;	
	Query Match 17.5%; Score 30.6; DB 8; Length 1992;	
	Best Local Similarity 44.7%; Pred.No. 2.5;	
	Matches 38; Conservative 13; Mismatches 34; Indels 0; Gaps 0	
Oy	30 UGACGUGAAACGGUCGUATGCGACACAACUGGUGUAUUGACUUACUANA CA 89	
	: : : : : : : : : : : : : : : :	
Dd	732 TAATTGTGAACACGGTCCAATTCTGTAATAAACCTTGATGCCAATTACATCAGCCTTGA 791	
Oy	90 AGAGCGCGUCGUGCAAGCCCAAAA 114	
Dd	792 TGATGGCGCAGATCATCGGAACA 816	
RESULT 12		
ID	ADA30421 standard; DNA; 2022 BP.	
XX		
AC	ADA30421;	
XX		
DT	20-NOV-2003 (first entry)	
XX		
DE	DNA encoding Acinetobacter baumannii protein #1708.	


```
XX de; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
KM vaccine; plant biocontrol agent.
XX
XX Acinetobacter baumannii.
OS
XX US6562958-B1.
XX
XX 13-MAY-2003.
PD
XX 04-JUN-1999; 99US-00328352.
PF
XX 09-JUN-1998; 98US-0088701P.
PR
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX Breton G, Bush D;
PI
XX WPI; 2003-576092/54.
XX
XX P-PSDB; ADA34547.
DR
XX
XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
XX Example; SEQ ID NO 1708; 328pp; English.
PS
XX
XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents DNA encoding an A. baumannii
CC protein.
CC
XX
XX Sequence 2022 BP; 648 A; 407 C; 410 G; 557 T; 0 U; 0 Other;
SQ
Query Match 17.5%; Score 30.6; DB 9; Length 2022;
Best Local Similarity 44.7%; Pred. No. 2.5;
Matches 38; Conservative 13; Mismatches 34; Indels 0; Gaps 0;
QY 30 UGAGUGAAAACGUGGUAUUGCUACAACACUGGUAUUGCUUACUUAUACA 89
DB 759 TAAATATGCAACGCGTCCAAATTGCTTAAAGCTTGATGCCAATTACCATCAGCTTGA 818
QY 90 AGAGCGCGUGGUGAGCCCAAAA 114
DB 819 TGATGCGCCAGATCATGCGAACA 843
RESULT 13
ABL28616/c
ID ABL28616 standard; DNA; 22188 BP.
XX
XX ABL28616;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 37321.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
PF
```

```
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 37321; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-
CC AB572072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 22188 BP; 6664 A; 5376 C; 4559 G; 5589 T; 0 U; 0 Other;
SQ
Query Match 17.5%; Score 30.6; DB 4; Length 22188;
Best Local Similarity 38.6%; Pred. No. 6.2;
Matches 39; Conservative 18; Mismatches 44; Indels 0; Gaps 0;
QY 46 CGUATUGCUACAACACUGGUAUUGCUUACUUAUACAAGACCGCGUGGAG 105
DB 19567 CGTAAATGCAACCAATATGGATCGCTATATGATTTTATTCATGCTGATGCTTCA 19508
QY 106 CCCACAAAAGUCUAGUAGUCACAGAGAGCAUAGCGUAG 146
DB 19507 CTTAGATCTGTTTTCGTAAGTGAAGCAGACATACGCGAG 19467
RESULT 14
ABZ80710
ID ABZ80710 standard; RNA; 188 BP.
XX
XX ABZ80710;
AC
XX
XX 15-OCT-2003 (first entry)
DT
XX
XX Cricket paralysis virus derived pseudoknot sequence.
DE
XX
XX Pseudoknot; secondary structure; cell-free protein synthesis; wheatgerm;
KM albumen; impurity; higher-order structure; intergenic region; IGR-IRIS;
XX internal ribosome entry site; ss.
XX
XX Cricket paralysis virus.
OS
XX
XX WO2003033719-A1.
XX
XX 24-APR-2003.
XX
XX 08-OCT-2002; 2002WO-JP010447.
XX
XX 17-OCT-2001; 2001JP-00319923.
XX
XX (NAAG-) NAT INST AGROBIOLOGICAL SCI.
XX
XX (WAKE-) WAKENYAKU KK.
XX
XX Nakashima N, Shibuya N, Nishikawa S;
XX
XX WPI; 2003-403230/38.
XX
```


	XX	Claim 1; Page 33; 39pp; Japanese.
PS	XX	The invention relates to to a cell-free protein synthesis system derived
CC	XX	from wheatgerm where there is substantial exclusion of wheatgerm embryo
CC	XX	albumen impurities. The novel system uses a sequence having a higher-
CC	XX	order RNA structure that promotes translation activity. The higher-order
CC	XX	sequence is preferably a "pseudoknot", especially derived from a range of
CC	XX	viruses (ABZ80707-ABZ80713). This sequence represents the "pseudoknot"
CC	XX	higher-order sequence from the Cricket paralysis virus. The sequence is
CC	XX	used in a construct which may also include an intergenic region and
CC	XX	internal ribosome entry site (IGR-IRES). The method is applicable in
CC	XX	producing useful proteins
SQ	XX	Sequence 188 BP, 56 A; 33 C; 36 G; 0 T; 63 U; 0 Other;
Query Match	17.3%;	Score 30.2; DB 8; Length 188;
Best Local Similarity	69.5%;	Pred. No. 1.4;
Matches	41; Conservative	0; Mismatches 18; Indels 0; Gaps 0
Oy	62	UUGGUAACUAAUUAGCUUUAUAUACAAGAGCGCCGUCGAGCCCAAGAAGCUGA 120
Db	73	UAGGUACCUAAUUAGCUUUAAGUCCAGAGAUCCUAGUGGACGCCCAAAUUAUCCAG 131
RESULT 15		
AAL50547	ID	AAL50547 standard; RNA, 200 BP.
AC	XX	AAL50547;
AD	XX	19-DEC-2002 (first entry)
DT	XX	Cricket paralysis virus RNA sequence.
DE	XX	
KM	XX	CripV-like virus; B8; higher-order structure; drug development;
RW	XX	drug production; translational activity-promoting function;
XX	XX	protein synthesis; structural analysis.
OS	XX	Cricket paralysis virus.
PN	XX	WO200261080-A1.
PD	XX	08-AUG-2002.
PF	XX	31-JAN-2001; 2001WO-JP000641.
PR	XX	25-JAN-2001; 2001JP-00016746.
PA	XX	(NNAAG-) NAT INST AGROBIOLOGICAL SCTI.
PI	XX	Nakashima N, Kanamori Y;
DR	XX	WPI; 2002-627482/67.
PT	XX	Translational activity-promoting higher-order structure of CripV-like
PT	XX	viruses for protein translation when suitably initiated, useful in
PT	XX	synthesis of proteins and polypeptides of foreign species for application
PT	XX	in drugs.
PS	XX	Claim 1; Fig 1-2; 38pp; Japanese.
CC	XX	The invention comprises seven RNA sequences (CripV-like viruses) which
CC	XX	have a higher-order structure that sustains translational activity-
CC	XX	promoting function. The RNA sequences of the invention are useful in the
CC	XX	synthesis of proteins and polypeptides for application in developing and
CC	XX	producing drugs. The RNA sequences of the invention are also useful in
CC	XX	basic research of protein synthesis and structural analysis by the gene
CC	XX	recombinant technique. The present nucleotide represents a Cricket
CC	XX	paralysis virus RNA sequence of the invention

```

XX      Sequence 200 BP; 60 A; 36 C; 37 G; 0 T; 67 U; 0 Other;
SQ
Query Match      17.3%; Score 30.2; DB 6; Length 200;
Best Local Similarity 69.5%; Pred. No. 1.5;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY      62  TTGGTGAACCAATTTTGAAGCCGTCGAGCCCAAAAGCTCAG 120
          |||||
Db       73  TAAAGTGAACCAATTTTGAAGCCGTCGAGCCCAAAAGCTCAG 131

```

Search completed: December 22, 2005, 08:40:43
Job time : 255.897 secs

```
Search completed: December 22, 2005, 08:40:43
Job time : 255.897 secs
```


TITLE
JOURNAL
COMMENT
Clifton, S., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagaris, R., Ronko, I., Kennedy, S., Maguire, L., Waterston, R. and Wilson, R.
Toxoplasma EST Project
Unpublished (2001)
Contact: Clifton, S.
Toxoplasma EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: toxo@wustl.wustl.edu
Contact David Sibley (toxos@borcim.wustl.edu) for further information relating to organism, libraries, or clone availability.
Seq primer: T7 from Gibco
High quality sequence stop: 588.
Location/Qualifiers

FEATURES
source
1..588
/organism="Toxoplasma gondii"
/mol_type="mRNA"
/db_xref="taxon:5811"
/clone="rgs27y32b05.y3"
/dev_stage="tachyzoite"
/lab_host="GC10 Competent Cells (PGC)"
/clone_lib="Tg COUG Tachyzoite cDNA Library"
/note="Vector: pDNR-LIB; Site 1: SfiI; Site 2: SfiI; The cDNA library was constructed by Kelian Tang, and Robert Cole at Washington University. Total RNA was converted to cDNA using the template-switching PCR method (Creator SMART cDNA, Clontech Inc.). First strand was PCR amplified using the same primer set and the fragments were digested with SfiI. The fragments were size selected, ligated into vector pDNR-LIB containing directional SfiI sites, and electroporated into GC10 competent cells. WARNING: the library contains a small percentage of cDNAs derived from the human host cells."

ORIGIN

Query Match 18.7%; Score 32.8; DB 7; Length 588;
Best Local Similarity 40.5%; Pred. No. 5.7;
Matches 60; Conservative 16; Mismatches 72; Indels 0; Gaps 0;

QY 11 GATCTGCGGCTGACGAGTCACTCTCGTGTGATGGGAGATGTACCAATCCGATCGT 70
DB 439 GATCTGCGGCTGACGAGTCACTCTCGTGTGATGGGAGATGTACCAATCCGATCGT 498
QY 71 UUUUUUAGCUUUAUUAAGAAGCGCGUGUGAGCCCAAAAGUCUUAUAGUCUACA 130
DB 499 TTTTGAGCAACAAGAGCGAGCGGCTGGGCAACCAAAATGGGTTCCGATGCAGTCA 558
QY 131 GGAGAGCAUACGCUAGGUGCGUUGACU 158
DB 559 GGTATCATACCATACGATGTGGTCTCT 586

RESULT 7
DR948423 856 bp mRNA linear EST 02-AUG-2005
LOCUS EST139962 Aquilegia cDNA library Aquilegia formosa x Aquilegia
DEFINITION pubescens cDNA clone COLRB13, mRNA sequence.
ACCESSION DR948423 GI:71717786
VERSION DR948423.1
KEYWORDS EST.
ORGANISM Aquilegia formosa x Aquilegia pubescens
SOURCE Aquilegia formosa x Aquilegia pubescens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales; Ranunculaceae; Aquilegia.
REFERENCE
1 (bases 1 to 856)
Hodges, S.A., Rensink, W., Buell, C.R., Borevitz, J., Kramer, E., Nordborg, M. and Tomkins, J.
Generation of ESTs from Aquilegia

JOURNAL

Unpublished (2005)
Other ESTs: EST139961
Contact: Scott Hodges
Department of Ecology, Evolution and Marine Biology
University of California, Santa Barbara
Santa Barbara, CA 93106, USA
Tel: 805 893 7813
Fax: 805 893 4724
Email: hodges@lifesci.ucsb.edu
Seq primer: M13 Reverse.
Location/Qualifiers

FEATURES
source
1..856
/organism="Aquilegia formosa x Aquilegia pubescens"
/mol_type="mRNA"
/db_xref="taxon:338618"
/clone="COLRB13"
/tissue_type="mixed shoot and floral apical meristems, flower buds, leaves and roots"
/lab_host="DH10B T1 (T1 and T5 phage resistance)"
/clone_lib="Aquilegia cDNA library"
/note="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI; F2, F3, and F4 lines of Aquilegia formosa x A. pubescens were grown from seed in greenhouses at UC Santa Barbara. From these plants three sets of tissue were collected: 1) Small flower buds (<10 mm) and very young inflorescences (71 & 29% by weight respectively). 2) Medium (7-20 mm) and large (at or near anthesis) flower buds (65 & 35% by weight respectively) and 3) Shoot apical meristems. A fourth set of tissue was collected from plants of A. formosa. These plants were grown from seed in sand and at approximately 1 month root tissue and leaf tissue of various developmental stages were collected (84 & 16% by weight respectively). Total RNA was extracted from each set of tissue and pooled in the following proportions: 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled total RNA, mRNA was extracted and enriched for full-length messages and then normalized with proprietary methods by Invitrogen."

ORIGIN

Query Match 18.6%; Score 32.6; DB 8; Length 856;
Best Local Similarity 42.5%; Pred. No. 7.2;
Matches 37; Conservative 16; Mismatches 34; Indels 0; Gaps 0;

QY 8 UGUAGUCUGCGCGAUAUAGUCUAGAAAGUCUGUAUUGCUACAACACUGUGU 67
DB 650 TGAATAACATCGCGGAAAAATCACTTGAAAAGGTCTTGCACTTCATCTTGGAG 709
QY 68 AGCUAUUAGCUUUAUUAAGAAGCG 94
DB 710 CGATTAATTGAAATTATTTATTTGGAGTGC 736

RESULT 8
CB460819 717 bp mRNA linear EST 26-MAR-2003
LOCUS 720727 MARC 680V Bos taurus cDNA 5', mRNA sequence.
ACCESSION CB460819
VERSION CB460819.1 GI:29267203
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE
1 (bases 1 to 717)
Smith, T.P.L., Roberts, A.J., Echterkamp, S.E., Chitko-McKown, C.G., Wray, J.E. and Keale, J.W.
A second set of bovine ESTs from pooled-tissue normalized libraries
Unpublished (2003)
Contact: Smith TP
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA

RESULT 11				
LOCUS	CO079367	787 bp	mRNA	linear
DEFINITION	GR_Ea42b19.f GR_Ea Gossypium raimondii cDNA clone GR_Ea42b19 5', mRNA sequence.			
ACCESSION	CO079367			
VERSION	CO079367.1	GI:48748848		
KEYWORDS	EST.			
SOURCE	Gossypium raimondii			
ORGANISM	Gossypium raimondii			
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.			
AUTHORS	1 (bases 1 to 787) Kim H., Yu Y., Kudrna D., Hatfield J., Stum D., Mueller C., and Udell J. A., Rapp R. A., Wendel J. F., Rao K., Soderlund C. and Wing R. A.			
TITLE	Global assembly of Cotton ESTs			
JOURNAL	Unpublished (2004)			
COMMENT	Contact: Rod A. Wing Arizona Genomics Institute The University of Arizona Forbes Building Room 303, Tucson, AZ, 85721-0036, USA Tel: 520 626 9595 Fax: 520 621 1259 Email: rwing@genome.arizona.edu Plate: 42 row: B column: 19. Location/Qualifiers			
FEATURES	1..787			
SOURCE	/organism="Gossypium raimondii" /mol_type="mRNA" /db_xref="caxon:29730" /clone="GR_Ea42b19" /issue_type="whole seedlings" /dev_stage="first true leaves" /lab_host="DH10B" /clone_id="GR_Ea" /note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2: EcoRI; library made by invitrogen from NotI-EV. Colonies scored; lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."			
ORIGIN				
Query Match	18.1%;	Score 31.6;	DB 7;	Length 787;
Best Local Similarity	39.2%;	Pred. No. 16;		
Matches	40;	Conservative 18;	Mismatches 44;	Indels 0;
QY	22 AATAAGCGTACGCGUGAAACGUGGCGUUGGUCACACACUGUGUAGCUAUNUAGCUU 81			
DB	651 AATATATCTCTGAGAAAGAGCCCTTTCCTACTGTAACAACAATCGTTAATTTGACTTG 710			
QY	82 ACUAAUCCAGACGCCGUGCGUGAGCCACACAAAGUCUAGUA 123			
DB	711 TCACATCCAAACATGCTCCCAAGCCTTCCAAATGCACAAATA 752			
RESULT 12				
LOCUS	AJ812852/c	800 bp	mRNA	linear
DEFINITION	AJ812852 KN206 Bos sp. cDNA clone C000520520, mRNA sequence.			
ACCESSION	AJ812852			
VERSION	AJ812852.1	GI:51880328		
KEYWORDS	EST.			
SOURCE	Bos sp.			
ORGANISM	Bos sp.			
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.			
AUTHORS	1 (bases 1 to 800) McGuire K. and Glass, F. J.			
TITLE	The expanding role of microarrays in the investigation of			

JOURNAL COMMENT

macrophage responses to pathogens
Vec. Immunol. Immunopathol. 105 (3-4), 259-275 (2005)
Contact: McGuire K
Genomics and Genetics
Roslin Institute
Roslin, Midlothian, EH25 9PS, UNITED KINGDOM
Single pass sequencing. Bases called and trimmed with phred v0.02045.c. Vector identified by cross_match with the -minscore 20 and -mismatch 12 options. Vector:PBHscRiptII(SK+) R. Site 1:EcoRV(lost) R. Site 2:NotI Seg Primer: T7 Normalised library constructed from pooled monocytes from Bos taurus (Holstein) and Bos indicus (Shihwals) cattle subjected to various stimuli,
including infection with the protozoan parasite Theileria annulata.

FEATURES

source Location/Qualifiers
1..800 /organism="Bos sp."
/mol_type="mrna"
/db_xref="taxon:29061"
/clone="C0005205P20"
/tissue_type="blood"
/cell_type="bovine monocyte"
/clone_id="KN206"
/note="Vector: PBHscRiptII(SK+); Site 1: EcoRV(lost); Site 2: NotI; Normalised library constructed from pooled monocytes from Bos taurus (Holstein) and Bos indicus (Shihwals) cattle subjected to various stimuli, including infection with the protozoan parasite Theileria annulata"

ORIGIN

Query Match 18.1%; Score 31.6; DB 1; Length 800;
Best Local Similarity 36.3%; Pred. No. 16;
Matches 37; Conservative 21; Mismatches 44; Indels 0; Gaps 0;

QY 22 AUAAGUCGACCGUAAAGCUUGCCGUANUGCUACAACACUGUGAUUUAGUU 81
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 248 ATATATATATATGTGGCTAATTTCCTTTCTTGTGAATAATGCATACACTTAGCCTT 169
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 82 ACUAAUCAAGACGCCGUCUGCACGCCCAAGAAGUCUAGUA 123
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 188 AGTTGAAAGAGCTGTCTGTGCACACTAACATTTTACAGA 147
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 13

AJ816139/c 806 bp mRNA linear EST 12-MAY-2005

LOCUS AJ816139 KN206 Bos sp. cDNA clone C0005205P20, mRNA sequence.
DEFINITION AJ816139
ACCESSION AJ816139
VERSION AJ816139.1 GI:51883615
KEYWORDS EST.
SOURCE Bos sp.
ORGANISM Bos sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 806)
McGuire,K. and Glass,E.J.
The expanding role of microarrays in the investigation of macrophage responses to pathogens
Vec. Immunol. Immunopathol. 105 (3-4), 259-275 (2005)
Contact: McGuire K
Genomics and Genetics
Roslin Institute
Roslin, Midlothian, EH25 9PS, UNITED KINGDOM
Single pass sequencing. Bases called and trimmed with phred v0.02045.c. Vector identified by cross_match with the -minscore 20 and -mismatch 12 options. Vector:PBHscRiptII(SK+) R. Site 1:EcoRV(lost) R. Site 2:NotI Seg Primer: T7 Normalised library constructed from pooled monocytes from Bos taurus (Holstein) and Bos indicus (shihwals) cattle subjected to various stimuli,
including infection with the protozoan parasite Theileria annulata

REFERENCE AUTHORS TITLE JOURNAL COMMENT

FEATURES

source Location/Qualifiers
1..806 /organism="Bos sp."
/organism="Bos sp."

Query Match	18.1%	Score 31.6	DB 7	Length 807
Best Local Similarity	39.2%	Pred. No. 16		
Matches 40, Conservative	18	Mismatches 44	Indels 0	Gaps 0

Search completed: December 22, 2005, 15:21:11
Job time : 1834.54 secs

Search completed: December 22, 2005, 15:21:11
Job time : 1834.54 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 07:55:17 ; Search time 73.5994 Seconds
(without alignments)
4226.575 Million cell updates/sec

Title: US-10-088-750C-7

Perfect score: 175

Sequence: 1 aguuuguugugauucuuugcgc.....acuaucuuauuauugacuu 175

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1103057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1 COMB. seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB. seq.*
3: /cgn2_6/ptodata/1/ina/6 COMB. seq.*
4: /cgn2_6/ptodata/1/ina/6 COMB. seq.*
5: /cgn2_6/ptodata/1/ina/6 COMB. seq.*
6: /cgn2_6/ptodata/1/ina/6 COMB. seq.*
7: /cgn2_6/ptodata/1/ina/6 COMB. seq.*
8: /cgn2_6/ptodata/1/ina/6 COMB. seq.*
9: /cgn2_6/ptodata/1/ina/6 COMB. seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	30.6	17.5	1345	US-09-270-767-1072	Sequence 1072, Ap
2	30.6	17.5	1345	US-09-270-767-16354	Sequence 16354, A
3	30.6	17.5	2022	US-09-328-352-1708	Sequence 1708, Ap
4	28.8	16.5	89689	US-09-949-016-13089	Sequence 13089, A
5	28.6	16.3	314798	US-09-949-016-13539	Sequence 13539, A
6	28	16.0	192506	US-09-949-016-15830	Sequence 15830, A
7	27.8	15.9	119211	US-09-596-002-40	Sequence 40, Appl
8	27.6	15.8	601	US-09-949-016-14624	Sequence 14624, A
9	27.6	15.8	1374	US-09-134-000C-1149	Sequence 1149, Ap
10	27.4	15.7	700	US-09-270-767-4977	Sequence 4977, Ap
11	27.4	15.7	700	US-09-270-767-20259	Sequence 20259, A
12	27.4	15.7	16397	US-08-956-171E-205	Sequence 205, App
13	27.4	15.7	16397	US-08-781-986A-205	Sequence 205, App
14	27.2	15.5	6076	US-09-824-734-1	Sequence 1, Appl
15	27.2	15.4	601	US-09-949-016-196049	Sequence 196049, A
16	27	15.4	753	US-09-540-236-132	Sequence 132, App
17	27	15.4	1095	US-09-774-796A-6516	Sequence 6516, App
18	27	15.4	5796	US-09-366-715-4	Sequence 4, Appl
19	27	15.4	6224	US-09-774-528-91	Sequence 91, Appl
20	27	15.4	6224	US-10-120-988-91	Sequence 91, Appl
21	27	15.4	41594	US-09-949-016-17298	Sequence 17298, A
22	27	15.4	41594	US-09-536-059-1	Sequence 1, Appl
23	27	15.4	65792	US-09-596-002-31	Sequence 31, Appl
24	27	15.4	177669	US-09-949-016-13713	Sequence 13713, A

25	26.6	15.2	425	3	US-09-270-767-11107	Sequence 11107, A
26	26.6	15.2	590	3	US-09-533-559-3654	Sequence 3654, Ap
27	26.6	15.2	2865	3	US-09-270-767-14817	Sequence 14817, A
28	26.6	15.2	94987	3	US-09-949-016-12510	Sequence 12510, A
29	26.6	15.2	154600	3	US-09-949-016-14757	Sequence 14757, A
30	26.6	15.2	165651	3	US-09-949-016-13032	Sequence 13032, A
31	26.4	15.1	444	3	US-09-248-796A-6916	Sequence 6916, Ap
32	26.4	15.1	601	3	US-09-949-016-27382	Sequence 27382, A
33	26.4	15.1	601	3	US-09-949-016-152512	Sequence 152512, A
34	26.4	15.1	693	3	US-09-248-796A-195	Sequence 195, App
35	26.4	15.1	1539	3	US-08-277-031B-5	Sequence 5, Appl
36	26.4	15.1	1539	3	US-08-277-031B-6	Sequence 6, Appl
37	26.4	15.1	1539	3	US-08-277-031B-7	Sequence 7, Appl
38	26.4	15.1	1655	3	US-09-949-016-4529	Sequence 4529, Ap
39	26.4	15.1	2509	3	US-09-949-016-4283	Sequence 4283, Ap
40	26.4	15.1	2553	3	US-09-949-016-468	Sequence 468, App
41	26.4	15.1	2565	3	US-09-023-655-1052	Sequence 1052, Ap
42	26.4	15.1	3466	3	US-09-027-166-10	Sequence 10, Appl
43	26.4	15.1	6387	2	US-07-721-775A-1	Sequence 1, Appl
44	26.4	15.1	6387	2	US-08-339-658-1	Sequence 1, Appl
45	26.4	15.1	7496	3	US-09-949-016-16025	Sequence 16025, A

ALIGNMENTS

RESULT 1
US-09-270-767-1072
; Sequence 1072, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1072
; LENGTH: 1345
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1072

Query Match 17.5%; Score 30.6; DB 3; Length 1345;
Best Local Similarity 47.5%; Pred. No. 0.21; Mismatches 19; Indels 0; Gaps 0;
Matches 29; Conservative 13;

QY	60	ACUUGUAGCUUUAUCUUAACAGACGCCGCGGACGCCACAAAGUUA	119
DB	588	ATTAGGTTACTATTACTGTTACGATGCTTATGGACCCCATATATCCA	647
QY	120	G	120
DB	648	G	648

RESULT 2
US-09-270-767-16354
; Sequence 16354, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16354
; LENGTH: 1345
; TYPE: DNA
; ORGANISM: Drosophila melanogaster

US-09-270-767-16354

Query Match	17.5%	Score 30.6	DB 3	Length 1345
Best Local Similarity	47.5%	Pred. No. 0.21		
Matches 29	Conservative 13	Mismatches 19	Indels 0	Gaps 0

Dy 60 ACUGGUAUAGUUACUAUCAGCGGUGCAGCCCAAAAAGUUA 119
| : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 588 ATTAGGTAACTATTAGTTACTGTCAGATGCTATTGGCAGCCCCAATAATATCCA 647

QY	120 G	120
Db	648 G	648

RESULT 3
US-09-328-352-1708
; Sequence 1708, Application US/09328352

```

GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 1708
LENGTH: 2022
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-1708

```

Query Match	17.5%	Score 30.6;	DB 3;	Length 2022;
Best Local Similarity	44.7%;	Pred. No. 0.25;		
Matches	38;	Conservative	13;	Mismatches 34;
			Indels	0;
			Gaps	0

Oy 30 UGACGCGAAAA CGUUGCGUAUVUCUACAACACUUGGUUAGCUAUVUAGCUTUACUAAUCA 89
 : | : | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 759 TAATATGCAAA GCGTCCAAATTGCTATAAAAGCTTGATGCCAATTACCATCAACGGTTGA 818

QY	90	AGACGCCGUCGUCAGCCCAACAAA	114
DB	819	TGATGGCGCAGTCATGCGAACA	843

RESULT 4
US-09-949-016-13089/c

```

: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CL001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: fastseq for Windows Version 4.0

```

US-09-949-016-13089

Query Match	16.5%	Score 28.8	DB 3	length 89689
Best Local Similarity	33.3%	Pred. No. 7.1		
Matches 48	Conservative 24	Mismatches 72	Indels 0	Gaps 0

60
 1 AGGUGUGUGUAGUUCGCGAUAAAUUGCUACGUGUAAAAGUUGCGUAGUUCUACACA
 60
 59969 ATTGTAGTTCCTTTCTTTCTTTTATTCTACGCGAAATGTCTCTTCAAGCTATTACT 59910
 Db

[illegible]

Oy	121	AUACGUCACAGAGAGCAUACGCU	144
		: :	
Db	59849	AAATTCTTAGAACAGCTTAATCT	59826

RESULT 5
US-09-949-016-13539
; Sequence 13539, Application US/09949016

```

1  APPLICANT: VENTER, J. Craig et al.
2  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
3  TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
4  FILE REFERENCE: CLO001307
5  CURRENT APPLICATION NUMBER: US/09/949,016
6  CURRENT FILING DATE: 2000-04-14
7  PRIOR APPLICATION NUMBER: 60/241,755
8  PRIOR FILING DATE: 2000-10-20
9  PRIOR APPLICATION NUMBER: 60/237,768
10 PRIOR FILING DATE: 2000-10-03
11 PRIOR APPLICATION NUMBER: 60/231,498
12 PRIOR FILING DATE: 2000-09-08
13 NUMBER OF SEQ ID NOS: 207012
14 SOFTWARE: FastSeq for Windows Version 4.0
15 SEQ ID NO. 13539
16 LENGTH: 314798
17 TYPE: DNA
18 ORGANISM: Human
19 FEATURE:
20 NAME/KEY: misc_feature
21 LOCATION: (1)...(314798)
22 OTHER INFORMATION: n = A,T,C or G
23
24 US-09-949-016-13539

```

Query Match	16.3%;	Score 28.6;	DB 3;	Length 314798;
Best Local Similarity	43.1%;	Pred. No. 15;		
Matches 53;	Conservative 11;	Mismatches 59;	Indels 0;	Gaps 0;

QY 16 UGCGCGAUAACUGCUGACGUGAAAAAGUUCGUUUGGCACCAACUUUGUAGCUAAU 75
 ::: :::
DB 294015 TGTCTAGGGCGCTGTGGAAAGGGAATGTAGGECTCAGAAGTCCACACAGAGTTACTACTCG 2940

[illegible]

QY	136	GCA	138
Db	294135	GCA	294137

RESULT 6
US-09-949-016-15830
: Sequence 15830, Application US/09949016

```

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF

```



```
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949.016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15830
LENGTH: 192506
TYPE: DNA
ORGANISM: Human
US-09-949-016-15830

Query Match          16.0%; Score 28; DB 3; Length 192506;
Best Local Similarity 33.9%; Pred. No. 20;
Matches 25; Conservative 21; Mismatches 30; Indels 0; Gaps 0;

QY 12 AUCUUGCGCCGUAUAUGCUGACGUGAAAACGUGGUAUUGCUACACACUUGGUAUCU 71
Db 117314 ATTTCACAGATAATTGAGAACTGAATGTTTAAACAACTCTCATTTAAAG 117373
QY 72 AUUAGCUUUAACUAU 87
Db 117374 TTTCAGCTTGATACT 117389

RESULT 7
US-09-596-002-40/c
Sequence 40; Application US/09596002
Patent No. 6632636
GENERAL INFORMATION:
APPLICANT: Lagace, Robert, E.
APPLICANT: Patterson, Chandra
APPLICANT: Berg, Kim, L.
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
FILE REFERENCE: FM-0008-4 US
CURRENT APPLICATION NUMBER: US/09/596.002
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: 60/140,121
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PERL Program
SEQ ID NO 40
LENGTH: 119211
TYPE: DNA
ORGANISM: M. catarrhalis
FEATURE:
NAME/KEY: m1ac_feature
OTHER INFORMATION: incycle template ID No. 6632636 40
PUBLICATION INFORMATION:
US-09-596-002-40

Query Match          15.9%; Score 27.8; DB 3; Length 119211;
Best Local Similarity 33.8%; Pred. No. 20;
Matches 41; Conservative 15; Mismatches 47; Indels 0; Gaps 0;

QY 56 CAACACUUGUUAUCUUAUUAUUAUUAUAUAUAAGACGCGUGUGACGCCACAAAG 115
Db 72121 CAATTTTGGATATCTTAAGCTTTGGCGATCTGATCTGAATCTTACCAAAAG 72062
QY 116 UCUAGUAUCGUCACAGAGACUAUCGUAAGUCGCGUGUAUCU 158
Db 72061 ACAACCAAAACCAAAAGACAGATGATGATGCTTTTACT 72019

RESULT 8
US-09-949-016-144624
Sequence 144624; Application US/09949016
Patent No. 6812339
```

```
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYOMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949.016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 144624
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-144624

Query Match          15.8%; Score 27.6; DB 3; Length 601;
Best Local Similarity 33.8%; Pred. No. 21;
Matches 25; Conservative 20; Mismatches 29; Indels 0; Gaps 0;

QY 12 AUCUUGCGCCGUAUAUGCUGACGUGAAAACGUGGUAUUGCUACACACUUGGUAUCU 71
Db 528 ATTTCACAGATAATTGAGAACTGAATGTTTAAACAACTCTCATTTAAAG 587
QY 72 AUUAGCUUUAACUA 85
Db 588 TTTCAGCTTGATA 601

RESULT 9
US-09-134-000C-1149
Sequence 1149; Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134.000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1149
LENGTH: 1374
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-134-000C-1149

Query Match          15.8%; Score 27.6; DB 3; Length 1374;
Best Local Similarity 36.6%; Pred. No. 3.1;
Matches 30; Conservative 18; Mismatches 34; Indels 0; Gaps 0;

QY 1 AGUGUGUGUAGUUCUGCGGUAUAUGCUGACGUGAAAACGUGGUAUUGCUACACA 60
Db 394 ATTCCTGGAGATGATGCGTATGATTTCTGGAAGAGCGTTGACGATTTGCTGAAGA 453
QY 61 CUUGGUAUCUUAUUAACUUA 82
Db 454 ATTGTTACCAAGATGATTA 475

RESULT 10
US-09-270-767-4977
Sequence 4977; Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
```


APPLICANT: Homburger et al.
 TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 FILE REFERENCE: File Reference: 7326-094
 CURRENT APPLICATION NUMBER: US/09/270,767
 CURRENT FILING DATE: 1999-03-17
 NUMBER OF SEQ ID NOS: 62517
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 4977
 LENGTH: 700
 TYPE: DNA
 ORGANISM: Drosophila melanogaster
 US-09-270-767-4977

Query Match 15.7%; Score 27.4; DB 3; Length 700;
 Best Local Similarity 34.0%; Pred. No. 2.8; Indels 0; Gaps 0;
 Matches 48; Conservative 22; Mismatches 71; Indels 0; Gaps 0;
 QY 19 GCGAUAUAGCUGAAGCGUAAACGUGGUAUUGCUACACACUUGGUUAGCUUUAAGC 78
 DB 4 GCAAAATATATTGAACCTAAACCATATATCATTAATGTGTAACTTGAACCTTAAGC 63
 QY 79 UUUACUAUACAAGACCGCGUGUGACGCCACAAAAGUCUAGUACGUCACAGAGAGA 138
 DB 64 TATATTATACCGAAGCGCTAAGTGGCTCTTACATTACACATTACACACATCTTACT 123
 QY 139 UACGCUAGUGCGGUGUACUA 159
 DB 124 GAACTAAGTCTTAAGTTTGT 144

RESULT 11
 US-09-270-767-20259
 Sequence 20259, Application US/09270767
 Patent No. 6703491
 GENERAL INFORMATION:
 APPLICANT: Homburger et al.
 TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 FILE REFERENCE: File Reference: 7326-094
 CURRENT APPLICATION NUMBER: US/09/270,767
 CURRENT FILING DATE: 1999-03-17
 NUMBER OF SEQ ID NOS: 62517
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 20259
 LENGTH: 700
 TYPE: DNA
 ORGANISM: Drosophila melanogaster
 US-09-270-767-20259

Query Match 15.7%; Score 27.4; DB 3; Length 700;
 Best Local Similarity 34.0%; Pred. No. 2.8; Indels 0; Gaps 0;
 Matches 48; Conservative 22; Mismatches 71; Indels 0; Gaps 0;
 QY 19 GCGAUAUAGCUGAAGCGUAAACGUGGUAUUGCUACACACUUGGUUAGCUUUAAGC 78
 DB 4 GCAAAATATATTGAACCTAAACCATATATCATTAATGTGTAACTTGAACCTTAAGC 63
 QY 79 UUUACUAUACAAGACCGCGUGUGACGCCACAAAAGUCUAGUACGUCACAGAGAGA 138
 DB 64 TATATTATACCGAAGCGCTAAGTGGCTCTTACATTACACATTACACACATCTTACT 123
 QY 139 UACGCUAGUGCGGUGUACUA 159
 DB 124 GAACTAAGTCTTAAGTTTGT 144

RESULT 12
 US-08-956-171E-205/c
 Sequence 205, Application US/08956171E
 Patent No. 6593114
 GENERAL INFORMATION:
 APPLICANT: Charles Kunsch
 Gil H. Choi
 Patrick S. Dillon

Craig A. Rosen
 Steven C. Barash
 Michael R. Fannon
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 5256
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/956,171E
 FILING DATE: 20-Oct-1997
 CLASSIFICATION: <Unknown>
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 60/009,861
 FILING DATE: January 5, 1996
 APPLICATION NUMBER: 08/781,986
 FILING DATE: January 3, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Mark J. Hyman
 REGISTRATION NUMBER: 46,789
 REFERENCE/DOCKET NUMBER: PB248P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (240) 314-1224
 TELEFAX: (301) 309-8439
 INFORMATION FOR SEQ ID NO: 205:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 16397 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 205:
 US-08-956-171E-205

Query Match 15.7%; Score 27.4; DB 3; Length 16397;
 Best Local Similarity 42.3%; Pred. No. 11; Indels 0; Gaps 0;
 Matches 30; Conservative 14; Mismatches 27; Indels 0; Gaps 0;
 QY 22 AUAUAGCUGAAGCGUAAACGUGGUAUUGCUACACACUUGGUUAGCUUUAAGCUTU 81
 DB 10332 AAATTGTTGATGGTAAAGTTCGTATCGCTAAATACTGCGAAGAAATTAATCT 10273
 QY 82 ACUAUACAAGA 92
 DB 10272 AATAATTATA 10262

RESULT 13
 US-08-781-986A-205/c
 Sequence 205, Application US/08781986A
 Patent No. 6737248
 GENERAL INFORMATION:
 APPLICANT: Charles Kunsch
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 5255
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
 COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781.986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248BP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8512
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 205:
SEQUENCE CHARACTERISTICS:
LENGTH: 16397 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-205

Query Match 15.7%; Score 27.4; DB 3; Length 16397;
Best Local Similarity 42.3%; Pred. No. 11;
Matches 30; Conservative 14; Mismatches 27; Indels 0; Gaps 0;

QY 22 AUAAGCUGACGUAACGUGGUAUUGCACAACUGGUAUUGCUU 81

DB 10332 AAATTGTGTAAGTAAAGTCTGCTAATAAATCGCGAAGAAATTAATCT 10273

QY 82 ACUAAUCAGA 92
DB 10272 AATAATTATA 10262

RESULT 14

US-09-824-734-1

Sequence 1, Application US/09824734
Patent No. 6727408
GENERAL INFORMATION:
APPLICANT: ZHU, JIAN-RANG
APPLICANT: SHI, HUAZHONG
APPLICANT: ISHITANI, MANABU
APPLICANT: STEVENSON, BECKY
TITLE OF INVENTION: PROTEINS AND DNA RELATED TO SALT TOLERANCE IN PLANTS
FILE REFERENCE: 205644US20
CURRENT APPLICATION NUMBER: US/09/824,734
CURRENT FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: US 60/194,648
PRIOR FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 6076
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(226)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (330)..(380)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (459)..(550)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (631)..(708)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (804)..(961)

OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (1044)..(1145)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (1235)..(1352)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (1432)..(1484)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (1571)..(1636)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (1738)..(1782)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (1933)..(1985)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (2470)..(2654)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (2767)..(2811)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (2899)..(3006)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (3222)..(3453)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (3531)..(3830)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (4012)..(4109)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (4193)..(4324)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (4407)..(4625)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (4712)..(4915)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (5003)..(5273)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (5375)..(5673)
OTHER INFORMATION:
NAME/KEY: CDS
LOCATION: (5771)..(6073)
OTHER INFORMATION:
US-09-824-734-1

Query Match 15.5%; Score 27.2; DB 3; Length 6076;
Best Local Similarity 36.1%; Pred. No. 8.8;
Matches 26; Conservative 18; Mismatches 28; Indels 0; Gaps 0;

QY 2 GUGUGUGUGAUCUGGCGUAUAUUGCUGACGUGAAACGUGGUAUUGCACAAC 61

DB 1705 GTGTATATATAATTTTCCCGGTAATTGTGACAGGAAATGGATATATTGCAACAC 1764

QY 62 UUGGUUAGCUAU 73

DB 1765 TTGTATTATTAT 1776

RESULT 15
US-09-949-016-196049/c
Sequence 196049, Application US/09949016

Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 196049
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-196049

	Query Match	15.4%	Score 27;	DB 3;	Length 601;
	Best Local Similarity	42.9%;	Pred. NO. 3.7;	Mismatches 40;	Gaps 0;
	Matches	39;	Conservative 12;	Indels	
QY	55	ACAAACACUGGUAUUNUAGUCUUACUAUACAAGACCGCUGUCAGCCACAAA	114		
		: : :			
DB	441	ACCACCTGGGGGAATATTGATTTTCACAAACCGATTTTATGTAAAAACCTCAAAA	382		
		: : :			
QY	115	GUUCUGAUAUCGUCACAGAGACAUCGUA	145		
		: : :			
DB	381	CTCAAAATCCCCCTTAATAAACCATATGTA	351		

Search completed: December 22, 2005, 15:30:29
Job time : 75.5994 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 08:09:01 / Search time 407.348 Seconds
(without alignments) 3552.588 Million cell updates/sec

Title: US-10-088-750C-7

Perfect score: 175
Sequence: 1 agugugugugaugcugcgc.....acuaucuuauauugaccu 175

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	175	100.0	187	US-10-088-750B-7 Sequence 7, Appli
2	31.6	18.1	5853	US-10-257-737-1 Sequence 1, Appli
3	31.4	17.9	599	US-10-972-079-69060 Sequence 69060, A
4	31.4	17.9	2398	US-10-450-763-7997 Sequence 7997, Ap
5	31.2	17.8	201	US-10-088-750B-3 Sequence 3, Appli
6	30.6	17.5	1992	US-10-283-122A-8963 Sequence 8963, Ap
7	30.6	17.5	22188	US-11-087-143-40155 Sequence 40155, A
8	30.2	17.3	200	US-10-088-750B-4 Sequence 4, Appli
9	30.2	17.3	61718	US-10-741-600-17743 Sequence 17743, A
10	29.4	16.8	129297	US-10-737-082-89 Sequence 89, Appli
11	29.4	16.8	129297	US-10-765-790-89 Sequence 89, Appli
12	29.4	16.8	148062	US-10-367-094-93 Sequence 93, Appli
13	29.2	16.7	13229	US-09-815-242-3808 Sequence 3808, Ap
14	29.2	16.7	1371	US-09-815-242-6827 Sequence 6827, Ap
15	29	16.6	1905	US-10-741-849-6083 Sequence 6083, Ap
16	28.8	16.5	325	US-10-437-963-15045 Sequence 15045, A
17	28.8	16.5	571	US-10-027-632-94700 Sequence 94700, A
18	28.8	16.5	571	US-10-027-632-105546 Sequence 105546, A
19	28.8	16.5	571	US-10-027-632-94700 Sequence 94700, A
20	28.8	16.5	571	US-10-027-632-94700 Sequence 94700, A
21	28.6	16.3	383	US-09-770-791-138 Sequence 138, App
22	28.6	16.3	565	US-09-925-065A-306806 Sequence 306806, A
23	28.6	16.3	565	US-09-925-065A-306807 Sequence 306807, A

24	28.6	16.3	575	US-09-925-065A-27081 Sequence 27081, A
25	28.6	16.3	575	US-09-925-065A-27082 Sequence 27082, A
26	28.6	16.3	575	US-09-925-065A-27083 Sequence 27083, A
27	28.6	16.3	575	US-09-925-065A-27085 Sequence 27085, A
28	28.6	16.3	575	US-09-925-065A-27086 Sequence 27086, A
29	28.6	16.3	844	US-10-027-632-156001 Sequence 156001, A
30	28.6	16.3	844	US-10-027-632-156002 Sequence 156002, A
31	28.6	16.3	844	US-10-027-632-156003 Sequence 156003, A
32	28.6	16.3	844	US-10-027-632-156004 Sequence 156004, A
33	28.6	16.3	844	US-10-027-632-156005 Sequence 156005, A
34	28.6	16.3	844	US-10-027-632-156006 Sequence 156006, A
35	28.6	16.3	2056	US-10-437-963-39314 Sequence 39314, A
36	28.6	16.3	3001	US-10-027-632-115403 Sequence 115403, A
37	28.6	16.3	3001	US-10-027-632-115403 Sequence 115403, A
38	28.4	16.2	198	US-10-088-750B-5 Sequence 5, Appli
39	28.4	16.2	630	US-10-021-323-8870 Sequence 8870, Ap
40	28.4	16.2	688	US-10-021-323-17254 Sequence 17254, A
41	28.4	16.2	794	US-10-425-115-13207 Sequence 13207, A
42	28.4	16.2	22080	US-09-423-093A-4 Sequence 4, Appli
43	28.2	16.1	532	US-09-925-065A-799026 Sequence 799026, A
44	28.2	16.1	537	US-09-925-065A-783500 Sequence 783500, A
45	28.2	16.1	565	US-09-925-065A-306805 Sequence 306805, A

ALIGNMENTS

```
RESULT 1
US-10-088-750B-7
; Sequence 7, Application US/10088750B
; Publication No. US20040166486A1
; GENERAL INFORMATION:
; APPLICANT: NAKASHIMA, Nobuhiko
; APPLICANT: KANAMORI, Yasushi
; TITLE OF INVENTION: A Novel Higher-Order Structure With Promoting Translation
; FILE REFERENCE: 3190-015
; CURRENT APPLICATION NUMBER: US/10/088, 750B
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: JP P2001-016746
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/JP01/00641
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 7
; LENGTH: 187
; TYPE: RNA
; ORGANISM: Rhopalosiphum Padl Virus
US-10-088-750B-7

Query Match      100.0%; Score 175; DB 7; Length 187;
Best Local Similarity 100.0%; Pred. No. 7.3e-51;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AGUGUGUGUUAUCUGCGCGCAUAUAUCUGCAAGAAAGUGUGCUUAUCUGCAACA 60
DB      1 AGUGUGUGUUAUCUGCGCGCAUAUAUCUGCAAGAAAGUGUGCUUAUCUGCAACA 60
QY      61 CUGUGUGUUAUUAUAGCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 120
DB      61 CUGUGUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 120
QY      121 AUACGUCACAGAGAGCAUACGUCGUGCGUGUACUUAUUAUUAUUAUUAUUAUUAUUA 175
DB      121 AUACGUCACAGAGAGCAUACGUCGUGCGUGUACUUAUUAUUAUUAUUAUUAUUAUUA 175

RESULT 2
US-10-257-737-1/C
; Sequence 1, Application US/10257737
; Publication No. US20050164354A1
; GENERAL INFORMATION:
```


APPLICANT: INRA
APPLICANT: COMPAGNIE GERVAIS DANONE
APPLICANT: RHODIA CHIMIE
APPLICANT: ANBA, Jamila
APPLICANT: HAGEGE, Juliette
APPLICANT: CHAILLOU, St,phane
APPLICANT: BESANCON-YOSHPE, Iris
APPLICANT: FREMAUX, Christophe
APPLICANT: MENGAULT, Pierre
APPLICANT: REMAULT, Pierre
TITLE OF INVENTION: LACTIC ACID BACTERIA MUTANTS OVERPRODUCING EXOPOLYSACCHARIDES
FILE REFERENCE: 33339/24937
CURRENT APPLICATION NUMBER: US/10/257,737
CURRENT FILING DATE: 2002-10-16
PRIOR APPLICATION NUMBER: FR 0004971
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 5853
TYPE: DNA
ORGANISM: Streptococcus thermophilus
FEATURE:
NAME/KEY: CDS
LOCATION: (2169)..(3092)
US-10-257-737-1

Query Match 18.1%; Score 31.6; DB 9; Length 5853;
Best Local Similarity 40.7%; Pred. No. 5.9; Mismatches 54; Indels 0; Gaps 0;

Matches 48; Conservative 16; Mismatches 54; Indels 0; Gaps 0;

QY 6 UGUUGAUCUUCUGCGGAUAAUGCUGACGUGAAAAGUUGCGUAUUGCUACAACUUGG 65
DB 437 TGTCTCAAGTCCGGAACGAGTTCTGGCGCAAAATAAGCACTGCTGCAATGACACCATTA 378
QY 66 UUAUCUUAUUGCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 123
DB 377 ATTGATGACTTGATTCATCAATAGAAATCGTGTGATCCAAATTAATCAGAGATA 320

RESULT 3

US-10-972-079-69060
Sequence 69060, Application US/10972079
Publication No. US20050153317A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: ROSENFELD, David
APPLICANT: KERR, Richard
APPLICANT: BATES, Stephen
APPLICANT: HOLM, Tom
TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEER
TITLE OF INVENTION: LIVESTOCK
FILE REFERENCE: MM11110-2
CURRENT APPLICATION NUMBER: US/10/972,079
CURRENT FILING DATE: 2004-10-22
PRIOR APPLICATION NUMBER: US 60/514,333
PRIOR FILING DATE: 2003-10-24
NUMBER OF SEQ ID NOS: 96631
SOFTWARE: PatentIn version 3.1
SEQ ID NO 69060
LENGTH: 599
TYPE: DNA
ORGANISM: Chicken 19866894340881_1
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(16)
OTHER INFORMATION: n is any nucleotide
US-10-972-079-69060

Query Match 17.9%; Score 31.4; DB 9; Length 599;
Best Local Similarity 35.5%; Pred. No. 2.8; Mismatches 56; Indels 0; Gaps 0;
Matches 43; Conservative 22; Mismatches 56; Indels 0; Gaps 0;

QY 10 UGAUCUCCGCCGUAUAUUGCUGACUGAAAACUGCGUAUUGCUACAACUUGUUG 69
DB 50 TGTCTCAAGTCCGGAACGAGTTCTGGCGCAAAATAAGCACTGCTGCAATGACACCATTA 109
QY 70 CUUUAUUGCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 129
DB 110 CGATTTTCTATGTATGAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTA 169
QY 130 A 130
DB 170 A 170

RESULT 4

US-10-450-763-7997/C
Sequence 7997, Application US/10450763
Publication No. US20050196754A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 7997
LENGTH: 2398
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIMILAR
LOCATION: (114)..(965)
OTHER INFORMATION: 38% homologous to Dictyostelium discoideum TlpD, accession
US-10-450-763-7997

Query Match 17.9%; Score 31.4; DB 9; Length 2398;
Best Local Similarity 37.2%; Pred. No. 4.8; Mismatches 71; Indels 0; Gaps 0;
Matches 54; Conservative 20; Mismatches 71; Indels 0; Gaps 0;

QY 1 AGUGUGUGUAGUUCUGCGGAUAAUUGCUGACGUGAAAAGUUGCGUAUUGCUACAAC 60
DB 1674 AATGTCTGGAATTTTGTGGGAAATGAGGCCGAAAGCTGTGAAGAACAGTGGAGA 1615
QY 61 CUUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 120
DB 1614 GATGTGAGACATTTTCAAGCAAGTATGAGACATTCATGCAACCAACCAAGGAGGG 1555
QY 121 AUAUGCAGAGAGCAUAGCUUA 145
DB 1554 GCTGTAGTTGGGAGTTTCCGTTA 1530

RESULT 5

US-10-088-750B-3
Sequence 3, Application US/10088750B
Publication No. US20040166486A1
GENERAL INFORMATION:
APPLICANT: NAKASHIMA, Nobuhiko
APPLICANT: KANAMORI, Yasushi
TITLE OF INVENTION: A Novel Higher-Order Structure with Promoting Translation
TITLE OF INVENTION: Activity
FILE REFERENCE: 3150-015
CURRENT APPLICATION NUMBER: US/10/088,750B
CURRENT FILING DATE: 2002-03-20
PRIOR APPLICATION NUMBER: JP P2001-016746

;; PRIOR FILING DATE: 2001-01-25
;; PRIOR APPLICATION NUMBER: PCT/JP01/00641
;; PRIOR FILING DATE: 2001-01-31
;; NUMBER OF SEQ ID NOS: 12
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 3
;; LENGTH: 201
;; TYPE: RNA
;; ORGANISM: Drosophila C Virus
US-10-088-750B-3

Query Match 17.8%; Score 31.2; DB 7; Length 201;
Best Local Similarity 66.2%; Pred. No. 2.1;
Matches 45; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 53 CUACAACACUGGUGUAGCUUACUUAUACAAGCCGCGUGGACCCACAA 112
DB 65 CUUAUUAUUGGUUACUUAUUGUUCUGUUCAGAGUCCUAUUGGACCCCAUA 124

QY 113 AAGUCUAG 120
DB 125 AUUUCAG 132

RESULT 6

US-10-282-122A-8963
; Sequence 8963, Application US/10282122A
; Publication No. US20040029129A1

;; GENERAL INFORMATION:

;; APPLICANT: Wang, Liangsu
;; APPLICANT: Zamudio, Carlos
;; APPLICANT: Malone, Cheryl
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Kari
;; APPLICANT: Zykied, Judith
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John
;; APPLICANT: Carr, Grant
;; APPLICANT: Yamamoto, Robert
;; APPLICANT: Foreyth, R.

;; APPLICANT: Xu, H.

;; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;; FILE REFERENCE: ELITRA.034A

;; CURRENT APPLICATION NUMBER: US/10/282,122A

;; CURRENT FILING DATE: 2003-02-20

;; PRIOR APPLICATION NUMBER: 60/191,078

;; PRIOR FILING DATE: 2000-03-21

;; PRIOR APPLICATION NUMBER: 60/206,848

;; PRIOR FILING DATE: 2000-05-23

;; PRIOR APPLICATION NUMBER: 60/207,727

;; PRIOR FILING DATE: 2000-05-26

;; PRIOR APPLICATION NUMBER: 60/230,335

;; PRIOR FILING DATE: 2000-09-06

;; PRIOR APPLICATION NUMBER: 60/230,347

;; PRIOR FILING DATE: 2000-09-09

;; PRIOR APPLICATION NUMBER: 60/242,578

;; PRIOR FILING DATE: 2000-10-23

;; PRIOR APPLICATION NUMBER: 60/253,625

;; PRIOR FILING DATE: 2000-11-27

;; PRIOR APPLICATION NUMBER: 60/257,931

;; PRIOR FILING DATE: 2000-12-22

;; PRIOR APPLICATION NUMBER: 60/267,636

;; PRIOR FILING DATE: 2001-02-09

;; PRIOR APPLICATION NUMBER: 60/269,308

;; PRIOR FILING DATE: 2001-02-16

;; Remaining Prior Application data removed - See File Wrapper or PALM.

;; NUMBER OF SEQ ID NOS: 78614
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 8963
;; LENGTH: 1992
;; TYPE: DNA
;; ORGANISM: Acinetobacter baumannii
US-10-282-122A-8963

Query Match 17.5%; Score 30.6; DB 7; Length 1992;
Best Local Similarity 44.7%; Pred. No. 8.6;
Matches 38; Conservative 13; Mismatches 34; Indels 0; Gaps 0;

QY 30 UGACGUGAAACGUGGCUUUGCUACAACACUGGUGUAGCUUACUUAUACAUA 89
DB 732 TAAATGCAAAAGCGTCAAAATGCTAAAGCTGTGTAATGCAATACATCAGCTTGA 791

QY 90 AGACCGCGUGGACGCCACAAA 114
DB 792 TGATGCGCAGATCATGCGAAACA 816

RESULT 7

US-11-097-143-40165/C
; Sequence 40165, Application US/11097143
; Publication No. US20050208558A1

;; GENERAL INFORMATION:

;; APPLICANT: Venter, J. Craig

;; APPLICANT: et al.

;; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID

;; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE

;; FILE REFERENCE: CLO00728

;; CURRENT APPLICATION NUMBER: US/11/097,143

;; CURRENT FILING DATE: 2005-04-04

;; PRIOR APPLICATION NUMBER: 60/157,832

;; PRIOR FILING DATE: 1999-10-05

;; PRIOR APPLICATION NUMBER: 60/160,191

;; PRIOR FILING DATE: 1999-10-19

;; PRIOR APPLICATION NUMBER: 60/161,932

;; PRIOR FILING DATE: 1999-10-28

;; PRIOR APPLICATION NUMBER: 60/164,769

;; PRIOR FILING DATE: 1999-11-12

;; PRIOR APPLICATION NUMBER: 60/173,383

;; PRIOR FILING DATE: 1999-12-28

;; PRIOR APPLICATION NUMBER: 60/175,693

;; PRIOR FILING DATE: 2000-01-12

;; PRIOR APPLICATION NUMBER: 60/184,831

;; PRIOR FILING DATE: 2000-02-24

;; PRIOR APPLICATION NUMBER: 60/191,637

;; PRIOR FILING DATE: 2000-03-23

;; NUMBER OF SEQ ID NOS: 43008

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 40165

;; LENGTH: 22188

;; TYPE: DNA

;; ORGANISM: DROSOPHILA

US-11-097-143-40165

Query Match 17.5%; Score 30.6; DB 10; Length 22188;
Best Local Similarity 38.6%; Pred. No. 22;
Matches 39; Conservative 18; Mismatches 44; Indels 0; Gaps 0;

QY 46 GGUUUGCUACAACUGGUGUAGCUUUAUUGCUUUAUACAAGCCGUGGUGAG 105
DB 19567 GGTAAATGCAACATTAATGATGCGCTATATGATTTTATTCATGCTGATGCTCAGG 19508

QY 106 CCCACAAAAGUCUAGUACGUCAGAGACCAUACGUGAG 146
DB 19507 CTTGATCTGTTTTCGTAAAGTGAAGCAGACATACGCGAG 19467

RESULT 8

US-10-088-750B-4
; Sequence 4, Application US/1008750B
; Publication No. US20040166486A1

;; GENERAL INFORMATION:

;; APPLICANT: NAKASHIMA, Nobuhiko

;; APPLICANT: KANAMORI, Yasushi

;; TITLE OF INVENTION: A Novel Higher-Order Structure With Promoting Translation

;; TITLE OF INVENTION: Activity


```

; FILE REFERENCE: 3190-015
; CURRENT APPLICATION NUMBER: US/10/088, 750B
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: JP P2001-016746
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/J001/00641
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 200
; TYPE: RNA
; ORGANISM: Cricket Paralysis Virus
; US-10-088-750B-4

```

```

Query Match      17.3%; Score 30.2; DB 7; Length 200;
Best Local Similarity 69.5%; Pred. No. 4.8;
Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```

```

QY      62  UUGGUAACUUAUUUAGCUUUAUUAUCAAAGCCGCGUGGACGCCCAAAAGUCUAG 120
Db      73  UAGGUUAGCUUAUUUAGCUUUAAGUCCAGAGUGCCUAGGCGAGCCCAUAUVCAG 131

```

```

RESULT 9
US-10-741-600-17743
; Sequence 17743, Application US/10741600
; Publication No. US20050026169A1

```

```

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 17743
; LENGTH: 61718
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: msc_feature
; LOCATION: (1)...(61718)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17743

```

```

Query Match      17.3%; Score 30.2; DB 8; Length 61718;
Best Local Similarity 34.7%; Pred. No. 46;
Matches 51; Conservative 23; Mismatches 73; Indels 0; Gaps 0;

```

```

QY      20  CGAUAUUGCUGAGCGAUAAGCGUGCGUAUUGCUACAACACUGGUAUUGCUAUUUAGCU 79
Db     22846 CCGGAGATGCGAGAGTGAABAAACATGTCATTTAACATTTGATTAAGGCTTTTATCT 22905
QY      80  UUACUAUACAAGACGCCGCGUGGACGCCCAAAAGUCUAGUAGUCACAGAGAGCAU 139
Db     22906 TTTCCTTACAGCTTAAAGCCTTGCCGTTTAAATTCCTTTCGTGTCACAGGGATCAA 22965
QY      140 ACGCUAGUGCGGUGUACUAUUCUUAU 166
Db     22966 GCAGCAGTTAACGCTGACGTTCCCTGT 22992

```

```

RESULT 10
US-10-737-082-89
; Sequence 89, Application US/10737082
; Publication No. US20050130170A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison

```

```

; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2032
; CURRENT APPLICATION NUMBER: US/10/737,082
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 89
; LENGTH: 129297
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-737-082-89

```

```

Query Match      16.8%; Score 29.4; DB 9; Length 129297;
Best Local Similarity 39.6%; Pred. No. 1.2e+02;
Matches 44; Conservative 16; Mismatches 51; Indels 0; Gaps 0;

```

```

QY      48  UAUUGCUACAACACUGGUAUUGCUUUAUUGCUUUAUUAUCAAAGCCGCGUGGACGCC 107
Db     121076 TATTCCTAAGCATCTGCTTCACATTATGCGGTAAAGACAAGAAAGAGGCTGCTTCCGCC 121135

```

```

QY      108  CACAAAGUCUUAUUAUAGUCUACAGAGAGCAUAGCGUAGGCGGUGUACU 158
Db     121136 CAGATPACACAGAGAGTTCCTCATTTCACGCGGTGACCTGACT 121186

```

```

RESULT 11
US-10-765-790-89
; Sequence 89, Application US/10765790
; Publication No. US20050130172A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2035
; CURRENT APPLICATION NUMBER: US/10/765,790
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 89
; LENGTH: 129297
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-765-790-89

```

```

Query Match      16.8%; Score 29.4; DB 9; Length 129297;
Best Local Similarity 39.6%; Pred. No. 1.2e+02;
Matches 44; Conservative 16; Mismatches 51; Indels 0; Gaps 0;

```

```

QY      48  UAUUGCUACAACACUGGUAUUGCUUUAUUGCUUUAUUAUCAAAGCCGCGUGGACGCC 107
Db     121076 TATTCCTAAGCATCTGCTTCACATTATGCGGTAAAGACAAGAAAGGCTGCTTCCGCC 121135
QY      108  CACAAAGUCUUAUUAUAGUCUACAGAGAGCAUAGCGUAGGCGGUGUACU 158
Db     121136 CAGATPACACAGAGAGTTCCTCATTTCACGCGGTGACCTGACT 121186

```

```

RESULT 12
US-10-367-094-93
; Sequence 93, Application US/10367094
; Publication No. US20040170982A1

```



```

: GENERAL INFORMATION:
: APPLICANT: David W. Morris
: APPLICANT: Marc Malandro
: TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
: FILE REFERENCE: 529452001500
: CURRENT APPLICATION NUMBER: US/10/367,094
: CURRENT FILING DATE: 2003-02-14
: NUMBER OF SEQ ID NOS: 203
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 93
: LENGTH: 149062
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-367-094-93

```

	Query Match	Similarity	16.8%	Score 29.4	DB 7	Length 159062
Best Local	Similarity	39.6%	Pred. No. 1.3e+02			
Matches	44	Conservative	16	Matches	51	Indels 0; Gaps 0;
QY	48	U A U G C U C A C A C U G G U A G C U A U U U A G C U U A C U A A G A G C C G C U G A G C C	107			
Db	130101	T A T T C T T A A G C C A T C G T G T T C A C A T T A T A C C G T A A C A C A A G A A G G C T G C T T C C T G C C	130160			
QY	108	C A C A A A A G C U N A G A U C G U C A C A G A G A C A U A C G C U A G C G C U G A C U	158			
Db	130161	C A G A T A A C C A C A G A G A T T C C C A T T T C A C T T T C A G G T G T C A C C T C G A C T	130211			

```

RESULT 13
US-09-815-242-3908
/ Sequence 3908, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl L.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Zykkind, Judith W.
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ TITLE OF INVENTION: Prokaryotes
/ FILE REFERENCE: EDITRA.01A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-23
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 1410
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3908
/ LENGTH: 1329
/ TYPE: DNA
/ ORGANISM: Enterococcus faecalis
US-09-815-242-3908

```

Query Match	16.7%	Score 29.2;	DB 3;	Length 1329;
Best Local Similarity	36.6%;	Pred. No. 23;		
Matches 30;	Conservative 19;	Mismatches 33;	Indels 0;	Gaps 0;

OY 1 AGUCUGUGUCAUCUCGCCGAGUAUAUGCUGACGGUAAAAACCCUGCGUATUCGUACAACA 60

Db 391 ATTCTGGAGAGTGGGCGTCATTGATTCGTGGAAGACCGTTGACGATTGCTGAAGAA 450
 Oy 61 CTTGGTGAACUATUUAGCTUUA 82
 Db 451 ATTGGTATCCCGATTATGTTAA 472

```

RESULT 14
US-09-815-242-6827
Sequence 6827, Application US/09815242
Patent No US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6827
LENGTH: 1371
TYPE: DNA
ORGANISM: Enterococcus faecalis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1371)
US-09-815-242-6827

```

```

Query Match      16.7%; Score 29.2; DB 3, Length 1371;
Best Local Similarity 36.6%; Pred. No. 23,
Matches 30; Conservative 19; Mismatches 33; Indels 0; Gaps 0

QY      1 AGUGUGUGUGAUVUCGCGCAAAUUGCUGACGUGAAAAGUUGUCGUAACA 60
      : : : : : : : : : : : : : : : : : : : : : :
Db      391 ATTCTCGAGAGTANAGCGCTATGATTTCTGTCGAAAGAGCGTTACGATTCCTGAAGA 450
      : : : : : : : : : : : : : : : : : : : : : :

QY      61 CUUGGUAAGCUAUVUACCUUA 82
      : : : : : : : : : : : : : : : : : : : : : :
Db      451 ATTGTTACCCAGTTATGTAA 472
      : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-10-741-849-6083/c
; Sequence 6083, Application US/10741849
; Publication No. US20050019931A1
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Busey, Howard

```



```

; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 28773
; LENGTH: 1964
; TYPE: DNA
; ORGANISM: Bovine 19866880766482
US-10-750-185-28773
```

```

Query Match 15.8%; Score 27.6; DB 6; Length 1964;
Best Local Similarity 37.8%; Pred. No. 2;
Matches 37; Conservative 17; Mismatches 44; Indels 0; Gaps 0;
```

```

Qy 68 AGCUAUVUAGCUUUAUCUAAGAGCGGCGUGACGCCCAAAAGUCUAUAGUC 127
Db 885 AGACACTGTGTATATCTCAAAACACCTTGAAAGTAGATATTTCTATTTC 944
Qy 128 ACAGAGAGCAUACGCGUAGCGCGUGACUACUCCUA 165
Db 945 ACAGATGAGAACCTGATGTTGGGTTAAGTAACTTTA 982
```

```

RESULT 3
US-10-750-185-61874
; Sequence 61874, Application US/10750185
; Publication No. US20050260603A1
```

```

; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 61874
; LENGTH: 1787
; TYPE: DNA
; ORGANISM: Bovine 19866880589931
US-10-750-185-61874
```

```

Query Match 15.7%; Score 27.4; DB 6; Length 1787;
Best Local Similarity 34.2%; Pred. No. 2;
Matches 40; Conservative 21; Mismatches 56; Indels 0; Gaps 0;
```

```

Qy 50 UUGCUACAACACUGUGUACUUVUAGCUUUAUCUAAGAGCGGCGUGACGCCCA 109
Db 956 TTTCCCCCACTTATGTTGTTTCATCTTCAATCCCAATTTTAAATGTCGATCTCA 1015
Qy 110 CAAAAGUCUAGUACUCACAGAGAGCAUACGCUAGUGCGUGUACUACUACU 166
Db 1016 GCAGAACTGACGCGGTGAGATGTGGCCAGGCCCATGAGGCTTGTAATTTCTCAT 1072
```

```

RESULT 4
US-10-750-185-29226
; Sequence 29226, Application US/10750185
; Publication No. US20050260603A1
```

```

; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 29226
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Bovine 19866880832929
US-10-750-185-29226
```

```

Query Match 15.5%; Score 27.2; DB 6; Length 807;
Best Local Similarity 33.8%; Pred. No. 1;
Matches 27; Conservative 20; Mismatches 33; Indels 0; Gaps 0;
```

```

Qy 8 UGUAUCUUCGCGCAUUAUCUGAAGAAAGUGCGUUAUUGCUACAACACUUGU 67
Db 226 TTTAATTTAATGCTATATTTTAAAGACAATGTTTCACCTTTAAAGCATTTGGTG 285
Qy 68 AGCUAUVUAGCUUUAUCUAU 87
Db 286 AGAGTGCACCTTTCTATT 305
```

```

RESULT 5
US-10-750-185-49272/c
; Sequence 49272, Application US/10750185
; Publication No. US20050260603A1
```

```

; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 49272
; LENGTH: 1443
; TYPE: DNA
; ORGANISM: Bovine 19866880857844
US-10-750-185-49272
```

```

Query Match 15.0%; Score 26.2; DB 6; Length 1443;
Best Local Similarity 36.8%; Pred. No. 5;
Matches 35; Conservative 17; Mismatches 43; Indels 0; Gaps 0;
```

```

Qy 4 GUGUGUGAUCUUGCGCGAUAUAGCUGAAGAAAGUGCGUUAUUGCUACAACACU 63
Db 914 GTTTTGATTAATCTTAGATCTACCCAAATATATAGAACCAATATACATCAAACTT 855
Qy 64 GGUUAGCUUUAUAGCUUUAUCUAAGAGCGCGU 98
Db 854 TCTAAGTTATACAGTCTAGCTAATATAAAAGGCCAT 820
```


RESULT 6

US-11-117-187-193/c
Sequence 193, Application US/11117187
Publication No. US20050265650A1
GENERAL INFORMATION:
APPLICANT: PREUSS, DAPHNE
APPLICANT: COPENHAVER, GREGORY
TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
FILE REFERENCE: ARCD:309US
CURRENT APPLICATION NUMBER: US/11/117,187
CURRENT FILING DATE: 2005-04-28
PRIOR APPLICATION NUMBER: US/09/531,120
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/125,219
PRIOR FILING DATE: 1999-03-18
NUMBER OF SEQ ID NOS: 212
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 193
LENGTH: 103931
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-11-117-187-193

Query Match 15.0%; Score 26.2; DB 7; Length 103931;
Best Local Similarity 36.0%; Pred. No. 41;
Matches 40; Conservative 18; Mismatches 53; Indels 0; Gaps 0;

QY 63 UGUUAGCUUUVUACUUAUACAAGCCGUGACGCCCAAAAGUCUAGU 122

DB 49398 TAGCTCACCATTTAGTTCTTCTGCTATACGACTTCGGTATCCAGAGACCTTAGAA 49339

QY 123 ACUCACAGAGAGAUACGUGUGCGUGACUUAUUAUAGAC 173

DB 49338 ACGAGCCACTTGTCTTAAGAACCTTCTTGTAAATGATCTTATTCATGAC 49288

RESULT 7

US-10-793-626-1047/c
Sequence 1047, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PUJ480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1047
LENGTH: 1071
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1047

Query Match 14.9%; Score 26; DB 6; Length 1071;
Best Local Similarity 37.8%; Pred. No. 6.1;
Matches 34; Conservative 16; Mismatches 40; Indels 0; Gaps 0;

QY 34 GUGAAACGUGGCUUAGUCUACAACUCUGUAGCUUUAUAGCUUAUACAAGAC 93

DB 445 GAGATTGACACAAAGTTTGCTTAACAATTTTGTAGACATAGATCTGCTCATCATTAAC 386

QY 94 GCCGUGGCGACGCCCAAAAGUCUAGAA 123

DB 385 CAACTGCTGCTCAACAATGTATTAAATTA 356

RESULT 8

US-10-793-626-2307
Sequence 2307, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PUJ480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2307
LENGTH: 1467
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-10-793-626-2307

Query Match 14.9%; Score 26; DB 6; Length 1467;
Best Local Similarity 40.9%; Pred. No. 7;
Matches 27; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

QY 26 AUGCUGACGUGAAACGUGGCUUAGUCUACAACUCUGUAGCUUUAUAGCUUUAUACUA 85

DB 722 AAGCTGGGGTGATGATGATTAATTAATGATACAGCTCATGCTCATTTCTAAAGCGTTAATTA 781

QY 86 AUCAG 91

DB 782 ATCAAG 787

RESULT 9

US-10-793-626-3736/c
Sequence 3736, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PUJ480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3736
LENGTH: 2352
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3736

Query Match 14.9%; Score 26; DB 6; Length 2352;
Best Local Similarity 40.9%; Pred. No. 8.7;
Matches 27; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

QY 26 AUGCUGACGUGAAACGUGGCUUAGUCUACAACUCUGUAGCUUUAUAGCUUUAUACUA 85

DB 2317 AAGCTGGGGTGATGATGATTAATTAATGATACAGCTCATGCTCATTTCTAAAGCGTTAATTA 2258

QY 86 AUCAG 91

DB 2257 ATCAAG 2252

RESULT 10


```
US-10-793-626-4242
; Sequence 4242, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348008
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4242
; LENGTH: 2980
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4242
```

```
Query Match          14.9%; Score 26; DB 6; Length 2980;
Best Local Similarity 40.9%; Pred. No. 9.7; Mismatches 25; Indels 0; Gaps 0;
Matches 27; Conservative 14; Mismatches 25; Indels 0; Gaps 0;
```

```
QY      26 AUGCUGACGUAACGUGGCUAUGCUACAACUGGUUAGCUAUAUUGCUUACUA 85
DB      1338 AAGCTGGCGTAGATGATTAATTGATGATACAGCTCATGTCATTCTAAAGCGTTATTA 1397
QY      86 AUCAG 91
DB      1398 ATCAAG 1403
```

```
RESULT 11
US-10-793-626-3390/c
; Sequence 3390, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348008
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3390
; LENGTH: 3502
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3390
```

```
Query Match          14.9%; Score 26; DB 6; Length 3502;
Best Local Similarity 37.8%; Pred. No. 10; Mismatches 40; Indels 0; Gaps 0;
Matches 34; Conservative 16; Mismatches 40; Indels 0; Gaps 0;
```

```
QY      34 GUGAAAACGUGGCUAUGCUACAACUGGUUAGCUAUAUUGCUUACUAUACAAGC 93
DB      1215 GAGAATTTCAGAACAGTTTGGCTTAACAATTTTATAGCAATAGATCTGCTCATCATTAAC 1156
QY      94 GCCGUGUGACGCCCAACAAGUCUGAUA 123
DB      1155 CAAGTCGTGTGTCACAAATTTGTTAATTA 1126
```

```
RESULT 12
US-10-793-626-4013
```

```
; Sequence 4013, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348008
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4013
; LENGTH: 3840
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4013
```

```
Query Match          14.9%; Score 26; DB 6; Length 3840;
Best Local Similarity 40.9%; Pred. No. 11; Mismatches 25; Indels 0; Gaps 0;
Matches 27; Conservative 14; Mismatches 25; Indels 0; Gaps 0;
```

```
QY      26 AUGCUGACGUAACGUGGCUAUGCUACAACUGGUUAGCUAUAUUGCUUACUA 85
DB      2186 AAGCTGGCGTAGATGATTAATTGATGATACAGCTCATGTCATTCTAAAGCGTTATTA 2245
QY      86 AUCAG 91
DB      2246 ATCAAG 2251
```

```
RESULT 13
US-10-995-561-13254/c
; Sequence 13254, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: C0601559
; CURRENT APPLICATION NUMBER: US/10/995,561
; PRIOR FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13254
; LENGTH: 65931
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13254
```

```
Query Match          14.7%; Score 25.8; DB 6; Length 65931;
Best Local Similarity 31.9%; Pred. No. 47; Mismatches 68; Indels 0; Gaps 0;
Matches 43; Conservative 24; Mismatches 68; Indels 0; Gaps 0;
```

```
QY      1 AGUGUGUGUGAUCUGGCGGUAUAUGCUAGUGGAAAACGUGGCUAUAUUGCUUACUA 60
DB      62650 AGTGGCGAGCTGATCCATGTAATAAAGACAAGATTAAATTTTGGTTTAAATGATACA 62591
QY      61 CUGGUAUGCUAUAUUGCUUUAUAUAAGACCGCGUGUGACGCCCAACAAGUCUAG 120
DB      62590 AAGTGTGAAGTTTAAAGTTTCTAATAATTATTCAGTATTTATCCCTTTCAAGATGAGG 62531
QY      121 AUAAGUCACAGAGA 135
DB      62530 GTTAGATACACAGA 62516
```

```
RESULT 14
US-10-750-185-61359/c
```



```

Sequence 61359, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMT GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FAVIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 61359
LENGTH: 752
TYPE: DNA
ORGANISM: Bovine 19866881166419
US-10-750-185-61359

```

Query Match	14.5%	Score 25.4	DB 6	Length 752
Best Local Similarity	36.1%	Pred. No. 8.7		
Matches 30	Conservative 17	Mismatches 36	Indels 0	Gaps 0

Dy 34 GUGAAACGUGCGUAUUGCAACAUCUGGUAGUAUUUAGCUUUAUCAUACAGAC 93
| : | | : : : : | : | : | : : : : | : | : |
Db 573 GTGAATCTTTTCCCTCTAAGCTATCATNGTCAGCTTTTTTAATTGGCAAATAAAAAT 514

```

QY      94 GCCGUCGUGCAGCCCAACAAAGU 116
          |||: ||| |||||:
DB      513 CTTGACTTCCAAGGAAAAAAGT 491

```

RESULT 15

```

US-10-750-185-54508/C
Sequence 54508, Application US/10750185
Publication No. US2005026063A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 54508
LENGTH: 788
TYPE: DNA
ORGANISM: Bovine
US-10-750-185-54508 19866881713019

```

Query Match	14.5%	Score 25.4	DB 6	Length 788
Best Local Similarity	37.3%	Pred. No. 8.9		
Matches 31	Conservative 16	Mismatches 36	Indels 0	Gaps 0

[illegible]

82 ACUAAUCAGAGCGCGUGCA 104

Db 302 AGCCCAACGAGCTCCTGTCCA 280

Search completed: December 22, 2005, 16:41:20
Job time : 167.698 secs

This Page Blank (uspto)

This Page Blank (uspto)